CELEBRATING 50

Public Sector Economics

3/2026

MISLAV BRKIĆ: Government debt dynamics in the EU after the pandemic: the role of the snowball effect

EVA HORVAT, ZRINKA ORLOVIĆ ALTABAS and DAVOR ZORIČIĆ: Comparative analysis of pension funds' performance in selected countries: are there significant Sharpe ratio differences?

ROBERT J. SONORA and JOSIP TICA: Household behavioural response to health policy recommendations: evidence from the COVID-19 pandemic in Bosnia and Herzegovina, Croatia, Serbia and Slovenia

DUŠAN PAVLOVIĆ, ŽIVKO BOJOVIĆ and VOJIN ŠENK: Transparent budget: the idea of deep transparency

GUILLERMO PEÑA: Applying VAT to loan-deposit intermediation

Vol. 50, No. 3 | pp. 329-460
September 2026 | Zagreb

ISSN: 2459-8860
https://doi.org/10.3326/pse.50.3

Institute of Public Finance

CELEBRATING 50

Public Sector Economics

3/2026

Vol. 50, No. 3 | pp. 329-460 | September 2026 | Zagreb

TABLE OF CONTENTS

Articles

Government debt dynamics in the EU after the pandemic: the role of the snowball effect

MISLAV BRKIĆ, univ. spec. oec.*

Article**
JEL: E31, E58, E62, E65, H63
https://doi.org/10.3326/pse.50.3.1

* The author would like to thank Saša Ljepović, Ana Šabić and two anonymous referees for their very helpful comments and suggestions.
Views and opinions expressed in the paper are the author's own and do not necessarily reflect the policies or views of the Croatian National Bank.

** Received: April 10, 2026
Accepted: July 22, 2026

Mislav BRKIĆ
Croatian National Bank, Trg hrvatskih velikana 3, 10000 Zagreb, Croatia
e-mail: mislav.brkic@hnb.hr
ORCiD: 0000-0003-3872-4824

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

Abstract

The COVID-19 pandemic and the subsequent economic recovery had a strong impact on government debt dynamics in the EU. This paper focuses on fiscal developments in the period 2021-2023, when the economic crisis triggered by the pandemic had already passed. In particular, it aims to determine whether the rapid reduction in government debt ratios in that period was driven by member states' primary surpluses or a favourable snowball effect – the positive difference between member states' nominal growth rates and the average interest rates on their debt. The analysis clearly reveals that the snowball effect was the key factor behind the favourable trend in debt ratios. The paper also discusses the challenges and adverse shocks that the EU faced in the post-pandemic period, which put pressure on national budgets and thus contributed to the persistence of primary deficits in that period.

Keywords: pandemic, inflation, fiscal policy, government debt

1 INTRODUCTION

Fiscal policy played a key role in supporting the EU economy during the COVID-19 pandemic. When the pandemic broke out in early 2020 and the national authorities responded by introducing mobility restrictions to contain infections, some parts of the economy were effectively shut down. For that reason, the mobility restrictions were complemented by large-scale fiscal support packages, aimed at protecting businesses and households against a dramatic fall in income. A strong fiscal intervention at the national and EU levels was justified given that the outbreak of COVID-19 was a purely exogenous shock that hit entire sectors indiscriminately. The fiscal packages adopted by member states' governments included wage subsidies and tax deferrals or exemptions, which were granted to all businesses that had suffered a major drop in revenues. These measures were a lifeline for many businesses and their employees – without them, corporate bankruptcies and the unemployment rate would certainly have spiked in 2020. Monetary policy played a decisive role too, as it managed to quickly restore stability to key financial markets, including euro area government debt markets.

The heavy fiscal costs of the pandemic coupled with the severe economic contraction caused a sharp increase in government debt levels. The debt ratio for the EU as a whole rose by 12 percentage points in 2020, reaching 90% of GDP. While all member states saw their debt ratios increase, the pace at which this happened was uneven across the EU – ranging from less than 5 percentage points in Bulgaria and Sweden, to more than 20 percentage points in Greece, Spain and Italy. This unfavourable trend was reversed already in the first half of 2021. The gradual easing of the pandemic enabled national authorities to lift the containment measures, which set the stage for a swift economic rebound. Inflation spiked in both the EU and globally, given that aggregate supply could not respond sufficiently fast to satisfy the rapidly growing demand supported by massive stimulus measures. The fiscal positions of member states improved markedly in such a context. The reduction in debt ratios was particularly fast in the case of highly indebted member states – in three cases, the debt ratio fell by more than 40 percentage points in less than three years.

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

This paper explores the factors behind the rapid reduction in member states' debt ratios in the period 2021-2023. By applying the standard decomposition of annual changes in the debt ratio, it is determined whether the decline in debt ratios was driven by strong primary fiscal balances or a favourable snowball effect – the positive difference between nominal GDP growth rates and the effective interest rates on debt. The analysis clearly shows that the snowball effect made a much larger contribution than the primary balances – in only one case was the decline in the debt ratio driven mainly by the accumulation of primary surpluses. This finding is consistent with the fact that the most indebted member states experienced the fastest reduction in their debt ratios after the pandemic. In addition, the analysis reveals that elevated inflation – triggered by the swift post-pandemic economic recovery and Russia's war against Ukraine – had a key role in generating the favourable snowball effect.

This paper complements the vast literature on the fiscal implications of the COVID-19 pandemic by taking a closer look at fiscal developments in the EU in the post-pandemic period. Several studies have explored the ambitious fiscal and monetary policy measures implemented at the peak of the pandemic (e.g. Chang, 2021; Cottarelli, 2021; Bökemeier and Wolski, 2022; Benigno et al., 2022; Della Posta and Morroni, 2022; Heimberger, 2023). The overall policy response to the pandemic has been assessed as timely, forceful and highly effective. Cottarelli (2021) and Heimberger (2023) have noted that the policy response was much stronger compared to previous crisis episodes, especially the euro area sovereign debt crisis. The existing papers have documented the sharp deterioration in fiscal fundamentals during the pandemic, whereby large fiscal deficits and a severe drop in nominal GDP contributed jointly to the sharp rise in debt levels. However, there seem to be no previous papers that deal systematically with fiscal developments in the EU after the pandemic. This paper aims to fill this gap. The early post-pandemic period was peculiar from a policy perspective as it was marked by strong economic and geopolitical forces with opposite effects on public finances. In particular, while the swift post-pandemic rebound and high inflation helped reduce budget deficits and debt ratios, frequent external shocks put negative pressure on budget balances.

The paper also contributes to the relatively scarce literature on the role of the snowball effect in the EU's public finances. Checherita-Westphal and Domingues Semeano (2020) found that the interest rate-growth differential had been positive – thereby producing an unfavourable snowball effect – most of the time between the early 1980s and 2019. The exception was the latest part of the period when the interest rate-growth differential turned negative due to the ECB's expansionary monetary policy. Rawdanowicz et al. (2021) came to similar conclusions after observing government debt dynamics in OECD member countries, including several EU member states, in the period 1970-2020. In addition, Rawdanowicz et al. (2021) analysed several episodes of sharp reductions in debt ratios and found that the main driver of debt reductions were primary surpluses, rather than the (favourable) snowball effect. This paper complements these works by showing how the snowball effect gained strength in the post-pandemic period and became the key driver of the rapid reduction in debt ratios. In particular, it shows how a strong

post-pandemic economic rebound coupled with high inflation on the one hand, and a relatively long average maturity of member states' government debt on the other, created an environment where the interest rate-growth differential turned deeply negative. The most indebted countries such as Greece, Spain, Italy and Cyprus experienced the fastest reductions in their debt ratios as a result.

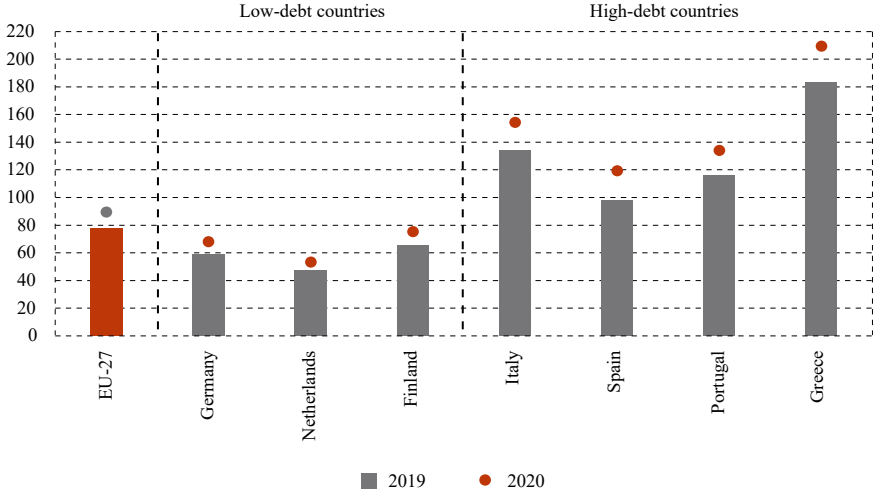
The rest of the paper is organized as follows. Chapter 2 provides a brief overview of the fiscal implications of the COVID-19 pandemic and the policy responses by individual member states and the EU collectively. Chapter 3 is the central part of the paper, which analyses the underlying drivers behind the rapid reduction in debt ratios after the pandemic, with a particular focus on the role of the snowball effect. Chapter 4 documents the fiscal policy challenges that member states encountered after the pandemic, which partly justify the persistence of primary deficits in that period. Chapter 5 presents the main conclusions of the paper.

2 THE FISCAL IMPACT OF THE PANDEMIC CRISIS IN THE EU

The COVID-19 pandemic was a perfect storm for public finances in the EU and around the world. Governments were prompted to significantly increase fiscal spending to cover pandemic-related costs at a time when fiscal revenues were falling sharply due to the lockdown-induced contraction in demand. The deterioration in fiscal indicators was further exacerbated by the steep fall in nominal GDP, which serves as the denominator when calculating the fiscal deficit and government debt ratios. Some EU member states were heavily indebted even before the outbreak of the pandemic. Market participants feared at the time that a further, pandemic-driven widening of budget deficits would make their government debt levels unsustainable, which was reflected in a sharp widening of sovereign bond yield spreads relative to Germany (Ortmans and Tripier, 2021).

FIGURE 1

Government debt of selected EU member states, 2019 and 2020 (percentage of GDP)



Source: Eurostat (2026a).

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

As shown in figure 1, although government debt levels increased significantly during the pandemic across the EU, there were notable differences in debt dynamics between member states. Southern countries with initial government debt levels of 100% of GDP or higher – notably Greece, Spain, Italy and Portugal – were the ones that experienced the most rapid rise in debt in 2020. In contrast, northern countries with more moderate initial levels of debt, such as Germany, Finland and the Netherlands, saw a much milder deterioration in fiscal positions during the pandemic. The wide variation in the fiscal impact was not surprising considering the differences in the severity of economic decline across countries. There are several papers analysing the heterogeneity in the member states' growth performance during the pandemic (e.g. Muggenthaler, Schroth and Sun, 2021; Ari, Atsebi and Palacios, 2022). Most of them point to the stringency of disease prevention measures and the sectoral composition of the economy as the two most important factors explaining the differences. Specifically, the southern countries experienced a relatively stronger contraction of output because they faced a particularly serious health crisis, which in turn prompted very strict social distancing, both government-imposed and voluntary. Moreover, as the economic weight of contact-intensive services is higher in the tourism-oriented southern countries, the sudden drop in demand for such services had a stronger impact on overall output than elsewhere in the EU.

The enormous scale of the pandemic shock required a swift and coordinated EU response, given that the fiscal capacity of some member states to handle the crisis on their own was insufficient. The ECB once again played a key role in mitigating the panic. In order to prevent a self-fulfilling run on fiscally vulnerable member states and ensure a smooth transmission of its monetary policy, the ECB announced in mid-March 2020 a large-scale government bond buying scheme – the Pandemic Emergency Purchase Programme (PEPP). This not only helped dispel investors' fears about debt sustainability, but also allowed euro area governments to meet their large financing needs at a low cost throughout the pandemic. In particular, the massive purchases by the Eurosystem in the secondary market ensured a high level of liquidity for euro area government bonds, thereby boosting investor confidence and reducing yield spreads[1]. The central banks of some non-euro area member states, including Poland, Romania and Croatia, also successfully implemented bond purchases, despite having no previous experience with such programmes (Arslan, Drehmann and Hofmann, 2020; CNB, 2020).

Fiscal policy responses had an equally strong positive impact on investor confidence. This particularly concerns the adoption of the generous EU-wide recovery plan funded by common borrowing (Corradin, Grimm and Schwaab, 2021). The

[1] Initially, the Eurosystem's massive bond purchases had a positive impact on market liquidity, as other investors found it easier to sell bonds at a favourable price. However, as the purchases continued, an increasing share of total outstanding government bonds became immobilised on the Eurosystem's balance sheet, implying that fewer bonds were freely floating in the market (Carrera de Souza and Hudepohl, 2024). The scarcity of bonds created difficulties in repo markets where government bonds are the main source of collateral. The Eurosystem alleviated this problem to some extent by conducting securities lending transactions, whereby government bonds purchased under the programme were lent back to financial institutions, typically against cash collateral.

PUBLIC SECTOR ECONOMICS
50 (3) 329–350 (2026)

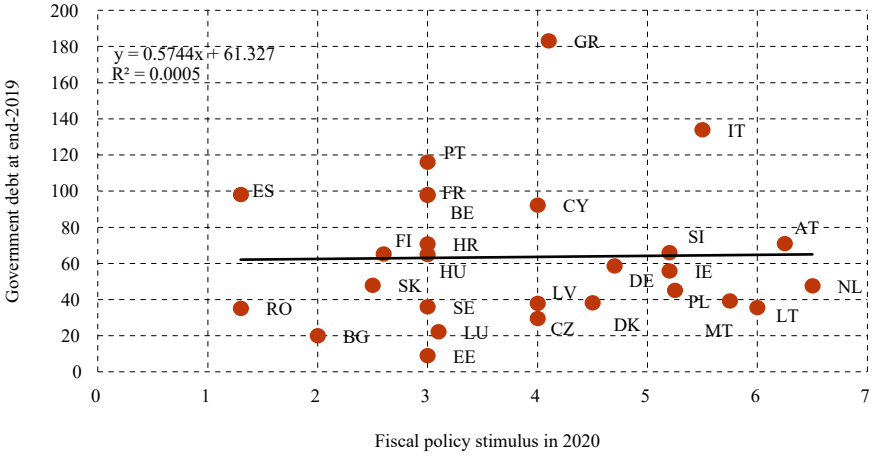
coordinated EU fiscal policy response started in March 2020 when the EU Council accepted the Commission's proposal for the activation of the general escape clause of the Stability and Growth Pact, allowing member states to temporarily deviate from fiscal rules. Then, in May 2020, the Council established the SURE facility, which helped finance member states' employment protection measures. Finally, in July 2020, the European Council established the NGEU instrument, with the Recovery and Resilience Facility at its core. The significance of the NGEU was twofold: it effectively transferred part of the fiscal costs of the pandemic from the strained national budgets to the EU level, and at the same time significantly enhanced the investment and growth potential of member states. Combined with a highly expansionary monetary policy, such a strong fiscal policy response enabled the EU economy to cope relatively well with the pandemic-induced crisis, without suffering lasting damage in terms of widespread corporate insolvencies, a loss of jobs and skills or a severe deterioration in banks' asset quality.

Despite the swift and ambitious EU response, member states' national budgets were heavily impacted by the crisis. The sharp widening of budget deficits during the pandemic was not only due to the operation of automatic stabilisers – falling tax revenues and rising unemployment benefits as economic activity declined. Discretionary fiscal easing measures played a major role too (European Commission, 2020). Specifically, shortly after the outbreak of the pandemic in March 2020, EU member states adopted ambitious fiscal stimulus packages to enable their economies to withstand a sharp contraction of activity triggered by the pandemic and the containment measures. In all member states, the stimulus packages included wage subsidies and tax and social security contributions deferrals for the affected companies. These two measures played a key role in supporting the liquidity of the corporate sector during the acute phase of the pandemic crisis, thus preventing mass bankruptcies and a surge in unemployment.

Interestingly, the size of the fiscal stimulus provided by individual member states during the pandemic crisis was not proportional to the fiscal space they had at that time. Specifically, the fiscal stimulus was large by historical standards in all member states, irrespective of their initial levels of government debt. As illustrated in figure 2, there was virtually no association between the member states' initial debt levels and the overall size of fiscal support measures adopted and implemented in the first year of the pandemic. For example, Italy and Greece adopted very large fiscal stimulus packages in early 2020, although their government debt ratios were the highest in the EU at the time, at above 130 and 180 percent of GDP, respectively. Supporting the economy with a massive counter-cyclical fiscal stimulus was justified in their case, as these two countries were about to suffer some of the worst output contractions in the EU. However, the strong fiscal expansion carried substantial risks as it implied that their already excessive government debt ratios would increase further significantly.

FIGURE 2

Fiscal policy stimulus and initial levels of government debt (percentage of GDP)*



*Note: *The fiscal policy stimulus refers to the estimated total size of pandemic-related support measures – including both revenue and expenditure side measures – adopted and implemented by national governments during 2020.*

Source: Eurostat (2026a), European Commission (2020).

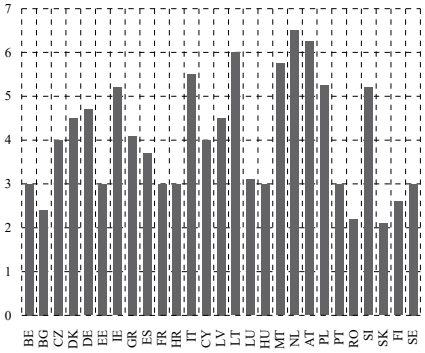
While during the pandemic crisis all EU member states relied heavily on fiscal policy measures to support their economies (figure 3a), fiscal policy had a more limited role during the previous major crisis – the global financial crisis of 2008-09 (figure 3b). In that particular episode, the overall size and even the direction of the fiscal response varied considerably across countries. In particular, only fiscally sound member states could afford to provide meaningful fiscal support to their economies. In contrast, countries that were considered fragile due to severe fiscal or external imbalances, or due to weaknesses in their banking systems, were not able to run a counter-cyclical fiscal policy. They even had to tighten fiscal policy in the midst of the recession to contain budget deficits and dispel investor concerns about the sustainability of their public finances. Despite fiscal consolidation efforts, many of them were eventually cut off from the financial markets and were forced to request emergency assistance from the IMF and the EU.[2]

[2] Between 2008 and 2013, eight EU member states requested financial assistance from the IMF and/or the EU. These were Latvia, Hungary, Romania, Greece, Ireland, Portugal, Spain and Cyprus.

FIGURE 3

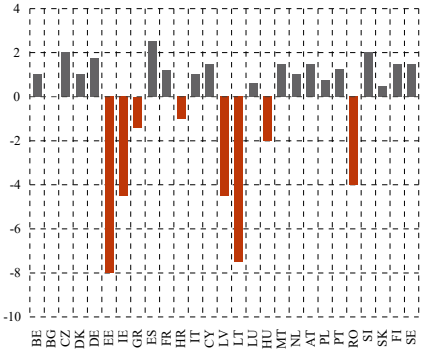
Fiscal policy stimulus by individual member states (percentage of GDP)*

a) Pandemic crisis (2020)



Source: European Commission (2020).

b) Global financial crisis (2009)



Source: European Commission (2009a) for Denmark and Italy, European Commission (2010) for Slovenia, IMF (2009a) for Greece, IMF (2009b) for Croatia, IMF (2009c) for Romania, and European Commission (2009b) for all other countries.

*Note: * The fiscal policy stimulus refers to the estimated total size of support measures – including both revenue and expenditure side measures – adopted and implemented by national governments during the respective year. If the amount of the fiscal stimulus is negative (red bar), it means that the country tightened fiscal policy during the crisis instead of providing fiscal stimulus.*

The reason why the fiscal policy response to the pandemic crisis was so strong and harmonised across the EU is related to the specific nature of that crisis. The COVID-19 pandemic was a global health emergency, which affected all member states irrespective of the state of their economies. In that regard, the pandemic crisis differed from an ordinary crisis which is typically driven by country-specific macroeconomic or financial vulnerabilities. Given that the pandemic crisis was essentially a liquidity crisis rather than a solvency crisis, there was a strong case for a coordinated EU policy response to cushion its impact. The decision by the ECB to launch the government bond purchase programme and the EU Council's decision to establish the NGEU were therefore justified. As mentioned above, these EU-wide policy interventions were crucial in reassuring investors and providing breathing space for the fiscally vulnerable member states, enabling them to carry out their own fiscal support packages. This largely explains why the relationship between the size of national fiscal support measures implemented in 2020 and initial levels of government debt was so weak (figure 2).

The case for a coordinated EU policy response is less clear if country-specific factors are at the root of the crisis. If individual member states are being penalised by financial markets for their excessive fiscal or financial imbalances, such a crisis should not be addressed by a common response such as a massive monetary expansion or EU-wide fiscal stimulus measures. Solvency issues of individual member states should be tackled primarily with tailor-made macroeconomic

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

adjustment programmes. The turmoil that the EU experienced after the outbreak of the 2008-09 global financial crisis is a notable example of such a crisis. Given that there was no major EU-wide fiscal policy response, the capacity of the vulnerable member states to provide fiscal support to their economies was severely constrained (figure 3b).

In conclusion, the outbreak of the pandemic was a major adverse shock for the EU that triggered an unprecedented economic policy response. Fiscal policy played a key role, alongside monetary policy, in shielding the economy and the financial system from liquidity disruptions brought about by the pandemic and the containment measures. Some member states implemented fiscal support packages far exceeding their objective fiscal capacities. This was made possible by the decisive measures taken by the ECB and the EU Council, which boosted investor confidence and thereby provided breathing space to national governments.

3 DRIVERS OF THE REDUCTION IN DEBT RATIOS AFTER THE PANDEMIC

The pandemic crisis was severe, but short-lived, and the economic recovery that followed was swift and robust. The recovery started in the third quarter of 2020, as soon as the gradual easing of the pandemic allowed the authorities to start lifting the containment measures. Given the short duration of the crisis and the ambitious monetary and fiscal policy response, the pandemic did not inflict permanent damage on the EU economy and the financial system. Although the economy shrank by 13% in real terms in the first half of 2020, it exceeded its pre-crisis level already in the third quarter of 2021. Moreover, thanks to generous wage subsidies, the labour market was largely intact – the average unemployment rate rose by only 0.4 percentage points, from 6.8% in 2019 to 7.2% in 2020. By 2022, it had fallen to a historical low of 6.2%.

The EU's recovery from the pandemic crisis was unusually strong from a historical perspective, particularly when compared to the slow and fragile recovery following the global financial crisis of 2008-09 and the euro area debt crisis of 2011-13. It took the EU economy almost seven years, until late 2014, to reach the pre-crisis level of output[3]. As noted above, that crisis was much more harmful for the EU because it triggered an abrupt unwinding of macroeconomic and financial imbalances that individual member states had accumulated up to that point (Lane, 2012; Gros, 2012).

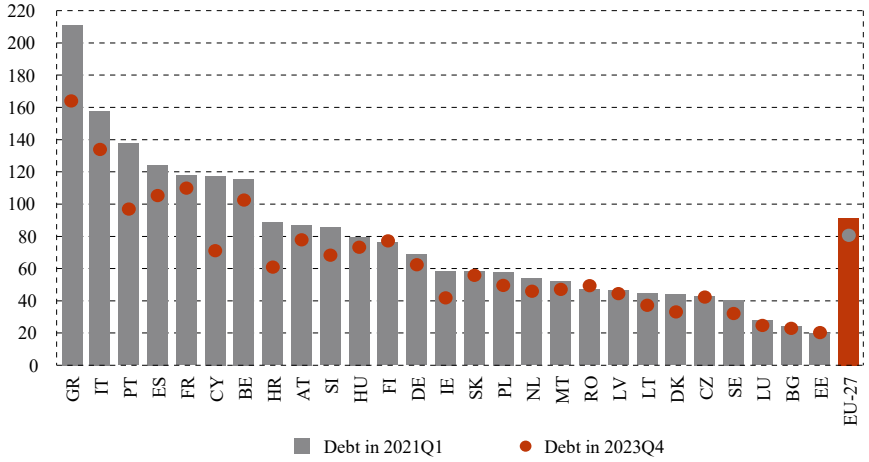
The robust recovery after the pandemic crisis enabled EU member states to repair their weakened fiscal positions relatively quickly. In particular, government debt ratios, which rose sharply in 2020 due to the recession and the high fiscal costs of the pandemic, started a steady decline the following year. The government debt ratio for the EU as a whole, after peaking at 91.5% of GDP in the first quarter of

[3] The highest pre-crisis level of quarterly real GDP was registered in the first quarter of 2008. This level of output was exceeded only in the fourth quarter of 2014, almost seven years later.

2021, declined steadily to 80.5% of GDP by the end of 2023. As shown in figure 4, almost all member states saw their government debt ratios improve in this period. Interestingly, the decline in the debt ratio was relatively more pronounced in countries with high initial levels of debt, namely Greece, Portugal, Cyprus, Italy and Spain. Indeed, as illustrated in figure 5, there was a strong statistical relationship between the level of government debt of member states in early 2021, and the cumulative decline in debt they experienced by the end of 2023.

FIGURE 4

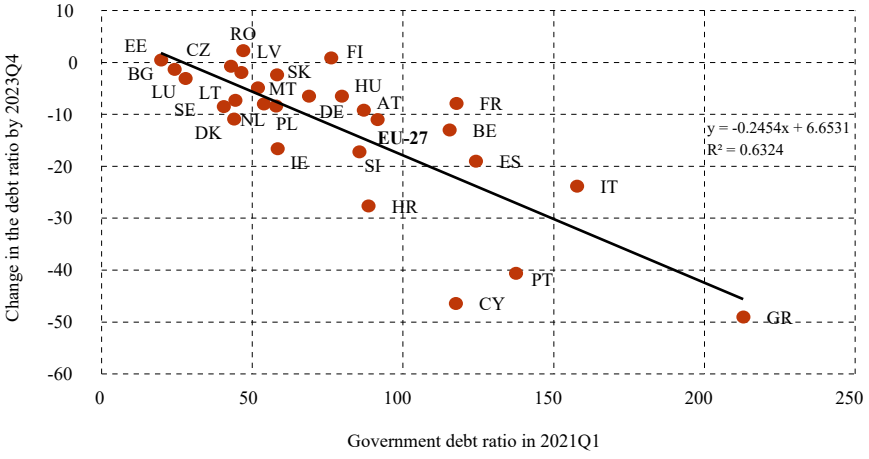
Government debt ratios: 2021Q1 vs. 2023Q4 (percentage of GDP)



Source: Eurostat (2026a).

FIGURE 5

Initial level of government debt and the cumulative decline in debt by 2023Q4 (percentage of GDP)



Source: Eurostat (2026a), author's calculations.

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

The pace at which government debt ratios of high-debt member states declined after the pandemic was unusually strong. In less than three years, between March 2021 and December 2023, the debt ratios of Greece, Portugal and Cyprus fell by 40-50 percentage points, and those of Spain and Italy by around 20-25 percentage points of GDP (figure 5). In contrast, the reduction in debt ratios was much more gradual during the recovery after the sovereign debt crisis (European Commission, 2018). More specifically, the cumulative reduction in the debt ratios of Greece, Spain, Italy and Portugal in the three-year period 2015-2017 was only 3% on average, compared to 29% in the period 2021-2023 (Eurostat, 2026a). It is therefore worth considering what specific factors were at play after the pandemic that enabled EU member states to reduce their elevated debt ratios so rapidly.

To investigate this, we need to decompose the annual changes in the government debt-to-GDP ratios into its underlying drivers. What follows is a derivation of the debt accumulation equation that will serve this purpose. The starting point is the identity that relates nominal government debt in the current year with nominal government debt in the previous year augmented by interest paid, and the primary government balance in the current year:

$$D_t = D_{t-1}(1+i_t) - PB_t \tag{1}$$

where D_t is nominal government debt at the end of year t, D_{t-1} is nominal government debt at the end of year $t-1$, i_t is the average weighted interest rate on government debt in year t, and PB_t is the primary government balance in year t. There are other factors that can also influence annual debt dynamics, such as acquisitions or sales of financial assets, but we are going to disregard them for this purpose.

To express identity (1) in terms of GDP, both sides of the equation have to be divided by Y_t:

$$D_t/Y_t = D_{t-1}(1+i_t)/Y_t - PB_t/Y_t \tag{2}$$

$$d_t = D_{t-1}(1+i_t)/Y_t - pb_t \tag{3}$$

where d_t is the government debt-to-GDP ratio in year t, and pb_t is the primary balance-to-GDP ratio in the same year. In order to have D_{t-1} also expressed in terms of GDP, we make use of the following identity:

$$Y_t = Y_{t-1}(1+g_t)(1+\pi_t) \tag{4}$$

where g_t is the real GDP growth rate, and π_t is the inflation rate.

After substituting Y_t with $Y_{t-1}(1+g_t)(1+\pi_t)$, we get:

$$d_t = D_{t-1}(1+i_t)/[Y_{t-1}(1+g_t)(1+\pi_t)] - pb_t \tag{5}$$

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

Finally, since $D_{t-1}/Y_{t-1} = d_{t-1}$, we arrive at the final equation:

$$d_t = d_{t-1}(1+i_t)/[(1+g_t)(1+\pi_t)] - pb_t \quad (6)$$

The debt accumulation equation derived above reveals that there are essentially two possible drivers of annual changes in the government debt ratio. The first driver is the ratio between the nominal interest rate on debt (represented by $(1+i_t)$) and the nominal GDP growth rate (represented by $(1+g_t)(1+\pi_t)$), also known as the snowball effect. The second driver is the primary fiscal balance (pb_t). The impact of these two drivers on the debt ratio is straightforward: the higher the nominal GDP growth rate relative to the interest rate on government debt, and the higher the primary balance, the lower the government debt-to-GDP ratio will be. Conversely, the higher the interest rate on debt relative to the nominal GDP growth rate, and the lower the primary balance, the higher the government debt ratio will be.

3.1 THE SNOWBALL EFFECT

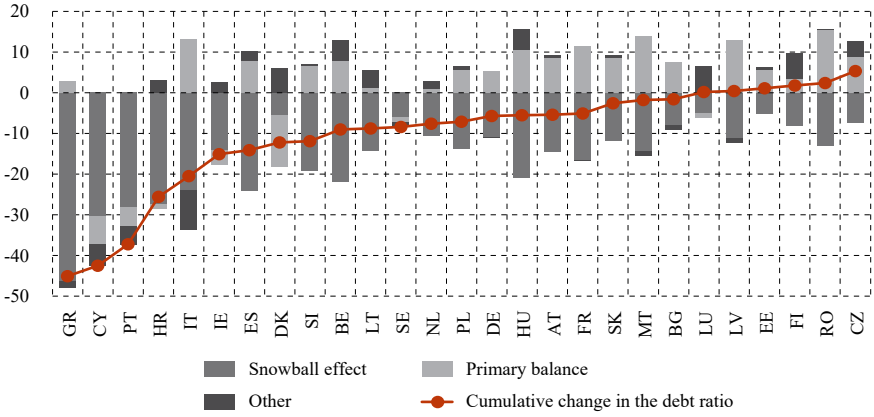
Figure 6 shows the decomposition of the cumulative change in government debt ratios in the period 2021-2023. This allows us to determine which of the two drivers – the snowball effect or the primary balance – played a more prominent role in the post-pandemic period. The figure clearly illustrates that the favourable snowball effect was the key driver of the fast reduction in debt ratios, especially in member states with elevated debt levels such as Greece, Cyprus, Portugal and Italy. Owing to the large positive difference between the nominal GDP growth rate and the average interest rate on their outstanding government debt, highly indebted countries were able to "grow out of their debt". More precisely, their government debt-to-GDP ratios declined simply because the denominator in the debt ratio – nominal GDP – expanded rapidly. The snowball effect had a favourable impact on the debt ratios of other member states as well, although this impact was generally weaker than in the case of highly indebted countries (figure 7). Among member states with a debt ratio below 90% of GDP, the snowball effect was particularly pronounced in the case of Croatia, given that it experienced the largest increase in nominal GDP in the entire EU.

The fact that highly indebted member states experienced a relatively stronger favourable snowball effect in the post-pandemic period than those with lower debt reflects the key feature of the snowball effect – it gains momentum as the debt ratio rises. Specifically, the difference between the nominal GDP growth rate and the interest rate on government debt produces larger absolute changes in the debt ratio when the initial debt ratio is high than when it is relatively low. In good times, when the growth-interest differential is positive, the (favourable) snowball effect enables highly indebted countries to reduce their debt ratios more rapidly. However, in bad times, when borrowing costs exceed the nominal GDP growth rate, the (unfavourable) snowball effect has an adverse impact on high-debt countries by further accelerating the rise in their debt ratios.

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

FIGURE 6

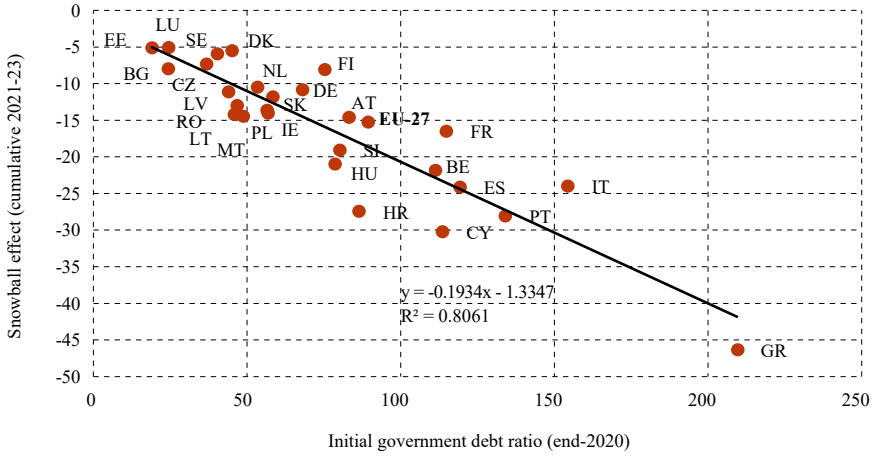
Decomposition of the cumulative change in government debt ratios in the period 2021-2023 (percentage points of GDP)



Source: Eurostat (2026a), author's calculations.

FIGURE 7

Initial levels of government debt and the favourable snowball effect in the period 2021-2023



Source: Eurostat (2026a), author's calculations.

As noted above, the period 2021-2023 was favourable for the public finances of EU member states, particularly for those that were heavily indebted. As an illustration, consider the case of Greece more closely. At the end of 2020, its government debt stood at 210% of GDP. The following year, Greece's nominal GDP expanded by about 10%, while the average interest rate on its debt remained low, at 1.3%. Due to the large positive growth-interest differential and the high initial level of debt, in 2021 alone the snowball effect produced a decline in the government debt ratio of 17 percentage points of GDP. If Greek debt at the end of 2020

had been 105% instead of 210% of GDP – the same growth-interest differential would have produced a snowball effect only half as large – 8.5 percentage points of GDP.

FIGURE 8

EU-27: Decomposition of the change in the government debt ratio (percentage points of GDP)

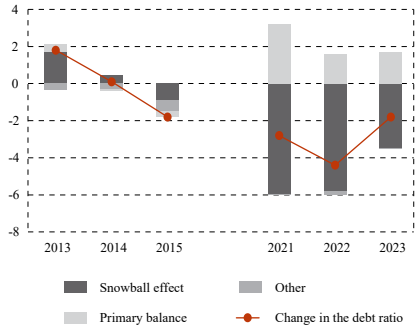
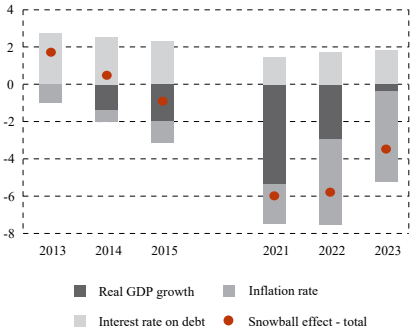


FIGURE 9

EU-27: Decomposition of the snowball effect (percentage points of GDP)



Source: Eurostat (2026a), author's calculations.

The snowball effect was not so supportive a decade earlier, when the EU was recovering from the sovereign debt crisis. Figure 8 compares the impact of the snowball effect on the debt ratio for the EU as a whole in the period 2013-2015 with its impact in the period 2021-2023. The difference between the two periods is evident. To illustrate what caused this large difference, figure 9 presents a decomposition of the snowball effect into its main drivers – real GDP growth, the inflation rate and the interest rate on debt. A couple of conclusions can be derived from these two figures.

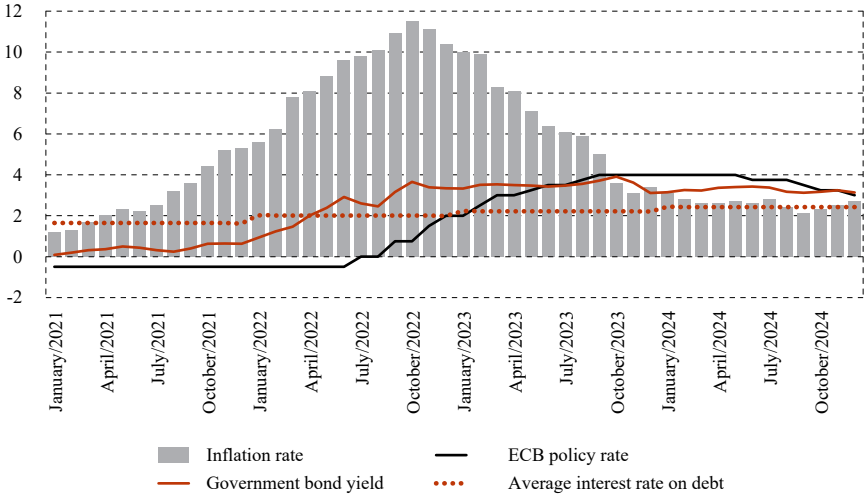
First, during the period 2013-2015, the combination of low growth and subdued inflation on the one side and moderate borrowing costs on the other created an unfavourable snowball effect, except in 2015 when it was mildly favourable. In such a setting, government debt ratios would have increased mechanically even if the primary fiscal balance had been zero. Therefore, accumulating primary surpluses was the only way national governments could bring down their debt ratios, but this was hard to achieve given the fragile economic conditions in that period. Second, in the period 2021-2023, the snowball effect had a favourable impact on debt ratios and this impact was very strong due to the robust post-pandemic economic growth, elevated inflation and relatively low borrowing costs. This enabled member states to achieve a significant reduction in their debt ratios despite running primary deficits.

Third, given that inflation supported the favourable snowball effect, particularly in 2022 and 2023, it is safe to say that a part of member states' government debt ratios was inflated away. Specifically, when inflation unexpectedly soared, this

was not followed by an equally sharp increase in member states' borrowing costs – government bond yields did rise, but much less than the inflation rate. There were two reasons for that. One reason was the stability of market participants' inflation expectations. Despite the sharp post-pandemic increase in the inflation rate, longer-term inflation expectations remained well anchored at the ECB's 2% target (ECB, 2024). Hence, while yields in government bond markets went up in 2022-2023 in response to the (expected) monetary policy tightening and higher uncertainty, they increased much less than the inflation rate (figure 10). The second reason was the relatively long average maturity of member states' government debt. Since only a fraction of the overall debt had to be refinanced in the high-inflation environment when yields in the bond market were higher, the pass-through of higher yields to the average cost of government debt was slow. In sum, as member states' average borrowing costs increased much less than the inflation rate, there is no doubt that high inflation contributed significantly to the favourable nominal growth-interest rate differential that brought down their debt ratios.

FIGURE 10

EU-27: Inflation rate, 10y bond yields and the ECB's policy rate (in percent)



Source: ECB (2026), Eurostat (2026b, 2026c, 2026d), author's calculations.

3.2 THE PRIMARY BALANCES

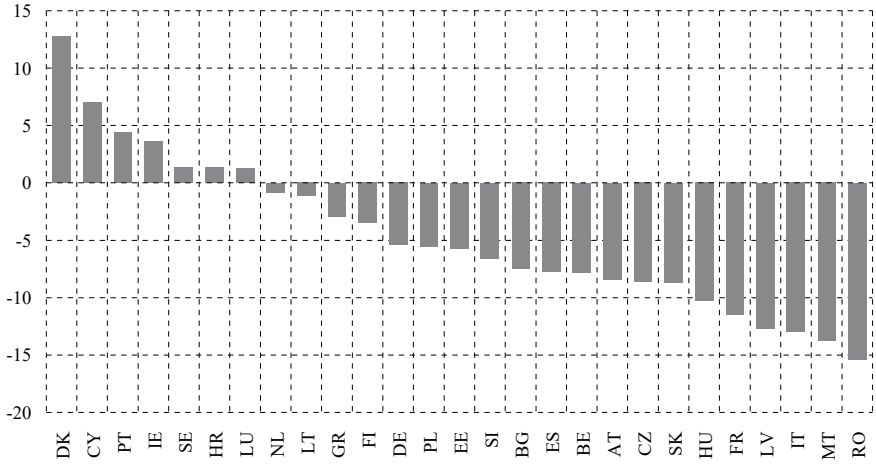
As explained above, the primary fiscal balance is the second driver determining the year-to-year changes in the government debt ratio. While the snowball effect put downward pressure on member states' debt ratios in the period 2021-2023, the primary balance in most cases had the opposite effect (figure 6). Specifically, only seven out of 27 member states recorded a positive primary balance, in cumulative terms, in the three-year period under review (figure 11). Denmark stood out in that regard with a particularly healthy fiscal balance. It was the only country where the decline in the debt-to-GDP ratio in the 2021-2023 period was driven mainly by primary surpluses, rather than the snowball effect. Among the countries which

recorded primary surpluses were also some highly indebted ones, namely Portugal and Cyprus.

In all other member states, the cumulative primary balance was negative, which means that it put upward pressure on the debt ratio. In several cases, the primary deficits were large enough to fully offset the favourable impact of the snowball effect, causing the government debt ratio to increase regardless of the positive growth-interest differential. Romania had the weakest fiscal performance among EU member states – its cumulative primary deficit in the period 2021-2023 amounted to more than 15% of GDP.

FIGURE 11

Primary fiscal balance, cumulative for 2021-2023 (percentage of GDP)



Source: Eurostat (2026a), author's calculations.

In sum, the analysis has shown that the rapid reduction in member states' government debt ratios after the pandemic was almost exclusively the result of the favourable snowball effect. Given that the vast majority of member states was accumulating primary deficits in this period, one could at first argue that fiscal policy did nothing to bring down government debt ratios. However, such a claim would not be entirely correct. Fiscal policy did help reduce debt ratios, but indirectly, by supporting domestic demand – and thus nominal GDP growth – during a challenging period marked by frequent adverse shocks. The next section touches upon these adverse shocks, as well as other challenges that partly explain the persistence of primary deficits in the post-pandemic period.

4 CHALLENGES FOR FISCAL POLICY IN THE POST-PANDEMIC PERIOD

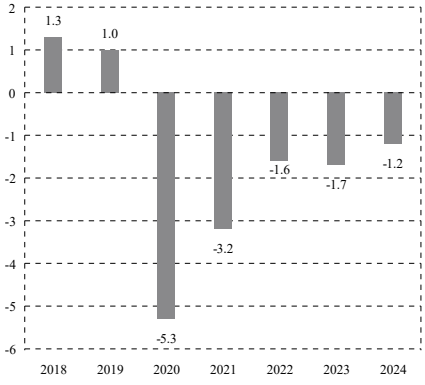
The fact that most member states ran primary deficits was a reflection of the delicate socio-economic and geopolitical environment that prevailed in the aftermath of the pandemic. At that time, there were three factors putting pressure on member states' national budgets. The first was the need to facilitate economic recovery after the

pandemic crisis. As mentioned earlier, the fiscal response to the pandemic was very ambitious, both the responses by individual member states and the collective response by the EU. The discretionary support measures introduced by member states in early 2020 remained sizeable throughout 2021, after which they were gradually phased out.

However, as soon as the pandemic-related fiscal spending slowed, the EU was confronted with another adverse shock – the sudden surge in inflation triggered by the energy price shock (European Commission, 2023a). This prompted member states to intervene with fiscal policy measures once again, this time to contain energy price increases and mitigate the adverse impact of inflation on households and businesses. Initially, the impact of inflation on the budget balances was positive, given that member states' fiscal revenues rose as soon as inflation soared, while it took some time for expenditures to catch up (European Commission, 2023b). Due to the gradual withdrawal of the pandemic-related measures and the positive initial impact of inflation on the budget, the primary balance for the EU as a whole improved in 2022 relative to the previous year (figure 12). However, in 2023, the primary balance deteriorated again slightly, as spending on pensions and public sector wages accelerated to compensate households for the loss of purchasing power caused by inflation. Figure 13 confirms that social benefits – which include pensions – and compensation of employees contributed significantly to the increase in government expenditures in 2023 and 2024. An analysis published by the ECB found that the high inflation episode had, in general, a negative impact on member states' budget balances, although it supported the reduction in the government debt ratio through the denominator effect (Bańkowski et al., 2023).

Figure 12

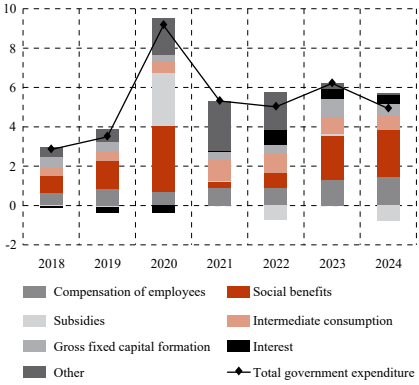
EU-27: Primary fiscal balance (percentage of GDP)



Source: Eurostat (2026a), author's calculations.

Figure 13

EU-27: Contributions to government expenditure growth (percentage points)



Source: Eurostat (2026b), author's calculations.

The sudden surge in inflation occurred in the context of a sharp increase in geopolitical risk, which in itself had budgetary implications. Specifically, Russia's unprovoked aggression against Ukraine in early 2022 highlighted two major sources of vulnerability for the EU – its excessive reliance on Russian energy supplies and

insufficient defence capabilities. The EU responded swiftly to enhance its energy security in such an environment. In May 2022, the European Commission proposed the RePowerEU plan – a strategic plan to reduce fossil fuel consumption, diversify away from Russia towards more reliable energy suppliers and expand its own capacities for renewable energy production (European Commission, 2022). Given that these efforts were aligned with the EU's green transition objective, funding for the RePowerEU Plan was secured under the existing Recovery and Resilience Facility, with additional support from EU cohesion policy funds (Schramm and Terranova, 2024). Hence, the steps to reduce dependence on Russian energy have had a limited impact on national budgets. More costly for member states' public finances have been efforts to enhance defence capabilities. Since the start of the war, most member states have markedly increased spending on defence, bringing total defence spending of the EU closer to the previous NATO target of 2% of GDP (figure 14). This trend has been particularly pronounced in countries in the geographical proximity to Russia and Ukraine, such as Poland and the Baltic States, where annual defence expenditures exceeded 3% of GDP by 2024 (figure 15). The EU is expected to continue raising spending on defence in the following decade, as member states make efforts to meet the new NATO target of 5% of GDP by 2035.[4]

FIGURE 14

EU-27: Total defence expenditure (EUR bn; percentage of GDP)

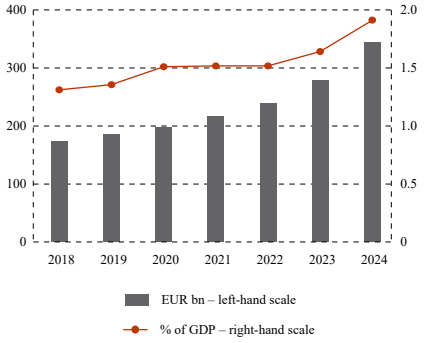
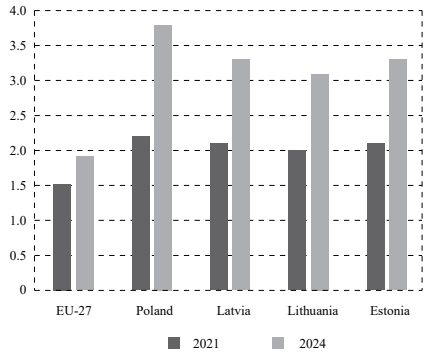


FIGURE 15

Total defence expenditure, EU-27 and selected countries (percentage of GDP)



Source: European Defence Agency (2026).

Taken together, these adverse factors put significant pressure on member states' public finances, given that a proactive use of fiscal policy was necessary to cushion the impact of the shocks. In such a delicate environment, it was objectively hard for member states to accumulate primary surpluses needed to bring down excessive debt ratios. Nevertheless, as explained earlier, member states managed to achieve a notable reduction in debt ratios thanks to resilient growth, high inflation and a relatively low average cost of borrowing.

[4] The new target was set at the NATO summit in Hague in June 2025. Member countries committed to raise core defence spending to at least 3.5% of GDP a year, as well as to allocate an additional 1.5% of GDP a year to defence-related activities, such as protecting critical infrastructure, improving civil preparedness and resilience and promoting innovation (NATO, 2025).

PUBLIC SECTOR ECONOMICS
50 (3) 329-350 (2026)

5 CONCLUSION

The outbreak of the pandemic in early 2020 marked the beginning of a very turbulent period for the EU. Several external shocks hit the EU one after another, generating large swings in economic growth and inflation rates. The large swings in growth and inflation rates were clearly mirrored in member states' government debt ratios. The pandemic crisis caused government debt ratios to soar, due to high budgetary costs of the pandemic and a sharp contraction in nominal GDP. In contrast, when the health situation improved and a robust economic recovery followed, accompanied by rising inflation, the fiscal positions of member states improved markedly. In particular, government debt ratios declined substantially, especially the debt ratios of highly indebted countries such as Greece, Portugal and Italy.

The analysis presented in this paper shows that the rapid reduction in debt ratios after the pandemic was almost exclusively the result of the favourable snowball effect – the large positive difference between member states' nominal GDP growth rates and the average interest rates on their debt. More specifically, the strong post-pandemic economic recovery coupled with high inflation rates generated unusually high nominal growth rates, above 8% on average for the EU as a whole in the period 2021-2023. Meanwhile, average interest rates on member states' outstanding debt rose much less, due to the long average maturity of debt and the fact that market yields did not increase as much as the inflation rate, owing to stable longer-term inflation expectations. In such a context, the nominal growth-interest differentials widened significantly, causing the government debt-to-GDP ratios to fall at a record pace.

The primary fiscal balances did not contribute much to the reduction in debt ratios. Only seven member states achieved a cumulative primary surplus in the 2021-2023 period. The remaining 20 member states recorded primary deficits, which were in some cases so large that they fully offset the positive impact of the snowball effect. The persistence of primary deficits after the pandemic can, however, be partly justified by the need to mitigate the adverse impact of the cost-of-living crisis and to boost defence spending after Russia's unprovoked aggression against Ukraine. Although the primary fiscal balances were negative in most countries, it would not be entirely correct to claim that the reduction in debt ratios was purely a result of luck, and that it was unrelated to fiscal policy. In fact, fiscal policy did contribute to such favourable trends, but indirectly, by supporting domestic demand – and thus nominal GDP growth – during the period marked by frequent adverse shocks.

Disclosure statement

The author has no conflicts of interest to declare.

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

No generative AI or AI-assisted tools were used in the preparation of this manuscript.

REFERENCES

1. Ari, A., Atsebi, J. M. B. and Palacios, M. D., 2022. Output Losses in Europe During COVID-19: What Role for Policies? *IMF Working Paper*, No. 22/130.
2. Arslan, Y., Drehmann, M. and Hofmann, B., 2020. Central Bank Bond Purchases in Emerging Market Economies. *BIS Bulletin*, No. 20.
3. Bańkowski, K. [et al.], 2023. The Effects of High Inflation on Public Finances in the Euro Area. *ECB Occassional Paper*, No. 332.
4. Benigno, P. [et al.], 2022. The European Monetary Policy Responses During the Pandemic Crisis. *Open Economies Review*, 33(4), pp. 657-675. https://doi.org/10.1007/s11079-022-09665-7
5. Bökemeier, B. and Wolski, M., 2022. This Time Is Different: Fiscal Response to the COVID-19 Pandemic Among EU Countries. *International Economics*, 172, pp. 217-226. https://doi.org/10.1016/j.inteco.2022.10.001
6. Carrera de Souza, T. and Hudepohl, T., 2024. Frictions in Scaling Up Central Bank Balance Sheet Policies: How Eurosystem Asset Purchases Impact the Repo Market. *Journal of Banking & Finance*, 158, 107037.
7. Chang, M., 2021. *Euro Area Fiscal Policies and Capacity in Post-Pandemic Times*. Paper prepared for the European Parliament Committee on Economic and Monetary Affairs. https://doi.org/10.2861/063465
8. Checherita-Westphal, C. and Domingues Semeano, J., 2020. Interest Rate-Growth Differentials on Government Debt: An Empirical Investigation for the Euro Area. *ECB Working Paper*, No. 2486.
9. CNB, 2020. *Macroeconomic Developments and Outlook*, No. 8. Zagreb: Croatian National Bank.
10. Corradin, S., Grimm, N. and Schwaab, B., 2021. Euro Area Sovereign Bond Risk Premia During the Covid-19 Pandemic. *ECB Working Paper*, No. 2561.
11. Cottarelli, C., 2021. *Euro Area Fiscal Policies and Capacity in Post-Pandemic Times*. Brussels: European Parliament.
12. Della Posta, P. and Morroni, M., 2022. The Credibility of Monetary Policy and the Fiscal Response to the Pandemic in the Eurozone. *Evolutionary and Institutional Economics Review*, 19, pp. 77-96. https://doi.org/10.1007/s40844-021-00226-0
13. ECB, 2024. *Annual Report 2023*. Frankfurt am Main: European Central Bank.
14. ECB, 2026. *Data Portal. ECB/Eurosystem policy and exchange rates*. Frankfurt am Main: European Central Bank.
15. European Commission, 2009a. *Public Finances in EMU 2009*. Brussels: European Commission.
16. European Commission, 2009b. *European Economic Forecast – Autumn 2009*. Brussels: European Commission.
17. European Commission, 2010. *European Economic Forecast – Spring 2010*. Brussels: European Commission.
18. European Commission, 2018. *Report on Public Finances in EMU 2017*. Brussels: European Commission.

19. European Commission, 2020. *European Economic Forecast – Autumn 2020*. Brussels: European Commission.
20. European Commission, 2022. REPowerEU: A plan to rapidly reduce dependence on Russian fossil fuels and fast forward the green transition. *Press release*, 18 May.
21. European Commission, 2023a. *Debt Sustainability Monitor 2022*. Brussels: European Commission.
22. European Commission, 2023b. *European Economic Forecast – Spring 2023*. Brussels: European Commission.
23. European Commission, 2024. *Report on Public Finances in EMU 2023*. Brussels: European Commission.
24. European Defence Agency, 2026. *EDA Defence Data*. Brussels: European Defence Agency.
25. Eurostat, 2026a. Government Finance Statistics (EDP and ESA 2010). *Government Deficit and Debt*. Brussels: Eurostat.
26. Eurostat, 2026b. Government finance statistics (EDP and ESA 2010). *Annual government finance statistics*. Brussels: Eurostat.
27. Eurostat, 2026c. Interest rates. *Maastricht criterion interest rates*. Brussels: Eurostat.
28. Eurostat, 2026d. Prices. *Harmonized index of consumer prices (HICP)*. Brussels: Eurostat.
29. Gros, D., 2012. Macroeconomic Imbalances in the Euro Area: Symptom or Cause of the Crisis. *CEPS Policy Brief*, No. 266.
30. Heimberger, P., 2023. This Time Truly Is Different: The Cyclical Behaviour of Fiscal Policy During the Covid-19 Crisis. *Journal of Macroeconomics*, 76, 103522. https://doi.org/10.1016/j.jmacro.2023.103522
31. IMF, 2009a. Greece: 2009 Article IV Consultation – Staff Report. *IMF Country Report*, No. 09/244.
32. IMF, 2009b. Republic of Croatia: 2009 Article IV Consultation – Staff Report. *IMF Country Report*, No. 09/185.
33. IMF, 2009c. Romania: Request for Stand-By Arrangement – Staff Report. *IMF Country Report*, No. 09/183.
34. Lane, P. R., 2012. The European Sovereign Debt Crisis. *Journal of Economic Perspectives*, 26(3), pp. 49-68.
35. Muggenthaler, P., Schroth, J. and Sun, Y., 2021. The Heterogeneous Economic Impact of the Pandemic Across Euro Area Countries. *ECB Economic Bulletin*, No. 5/2021.
36. NATO, 2025. *The Hague Summit Declaration*. Brussels: North Atlantic Treaty Organization.
37. Ortmans, A. and Tripier, F., 2021. COVID-induced Sovereign Risk in the Euro Area: When Did the ECB Stop the Spread? *European Economic Review*, 137, 103809. https://doi.org/10.1016/j.euroecorev.2021.103809

38. Rawdanowicz, L., Turban, S., Haas, J., Crowe, D. and Millot, V., 2021. Constraints and Demands on Public Finances: Considerations of Resilient Fiscal Policy. *OECD Economics Department Working Papers*, No. 1694.
39. Schramm, L. and Terranova, C., 2024. From NGEU to REPowerEU: Policy Steering and Budgetary Innovation in the EU. *Journal of European Integration*, 46(6), pp. 943-961. https://doi.org/10.1080/07036337.2024.2353791

Comparative analysis of pension funds' performance in selected countries: are there significant Sharpe ratio differences?

EVA HORVAT, Ph.D.*
ZRINKA ORLOVIĆ ALTABAS, Ph.D.*
DAVOR ZORIČIĆ, Ph.D.*

Article**
JEL: G23, G28
https://doi.org/10.3326/pse.50.3.2

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

** Received: April 29, 2026
Accepted: June 29, 2026

Eva HORVAT
Raiffeisen Mandatory and Voluntary Pension Funds Management Company, Petrinjska 59, 10000 Zagreb, Croatia
e-mail: eva.horvat@rmf.hr
ORCiD: 0000-0002-1000-2988

Zrinka ORLOVIĆ ALTABAS
Faculty of Economics and Business, University of Zagreb, Trg. J. F. Kennedy 6, 10000 Zagreb, Croatia
e-mail: zrinka.orlovic@efzg.unizg.hr
ORCiD: 0000-0003-2508-2658

Davor ZORIČIĆ
Faculty of Economics and Business, University of Zagreb, Trg. J. F. Kennedy 6, 10000 Zagreb, Croatia
e-mail: davor.zoricic@efzg.unizg.hr
ORCiD: 0000-0002-0206-3422

PUBLIC SECTOR ECONOMICS
50 (3) 351-372 (2026)

Abstract

The paper examines the risk-adjusted performance of pension funds for six countries based on country-level Sharpe ratios. Four of the countries are former transition economies (Croatia, Slovakia, Romania, and Bulgaria), which, together with Sweden, comprise the EU member states in the dataset. Chile is added as the sixth, non-EU member country to provide a more internationally balanced sample. Based on monthly data for the period from July 2015 to December 2024, the statistical significance of differences in Sharpe ratios is tested for the same risk-categories across countries and between risk-categories within each country. The Jobson-Korkie-Memmel, Opdyke and Ledoit-Wolf tests are performed. The results of empirical analysis suggest that the performance of pension funds in both A and B risk-categories is statistically significantly different only in instances when pairwise testing involves Croatia. When comparing risk-categories A and B within each country, a statistically significant difference is found only in Bulgaria.

Keywords: risk-adjusted performance, risk categories of pension funds, transition economies, statistical significance tests

1 INTRODUCTION

The performance of pension funds is always scrutinised both by the general public and in academic circles, given the numerous potential concerns which warrant research and discussion. Frequent research topics, therefore, range from analysing the problem of their underfunding or retirement income inadequacy risk (see, for instance, Giron et al., 2018) to (outrage by) fees being charged for their management (see, for instance, Dyck, Manoel and Morse, 2022). However, an even broader view can easily incorporate the pension system design which then includes the discussions about whether pension funds should be privately or publicly managed, whether there should be more than just one pension fund, and other critiques of the pension reform conducted around the world[1], most notably by ILO (2018). All of the above only adds to the general concern for the prudent management of pension funds' assets, which is why the industry is heavily regulated with investment constraints in place. Thus, pension funds cannot make fully discretionary decisions on how to allocate the assets under management. However, regulatory rules differ across countries and may also vary within a country, depending on the targeted exposure of different types of pension funds.

As presented by Brinson, Singer and Beebower (1991), a fund's return can generally be explained by three components: the investment strategy (asset allocation policy), the timing of investment decisions (market timing), and security selection. This approach is followed, for instance, by Blake, Lehmann and Timmermann (1999) and Andonov, Bauer and Cremers (2012), with Broeders and de Haan (2020) adding a fourth component, benchmark selection, to better describe a fund's realized return.

[1] As elaborated in ILO (2018) in the period from 1981 till 2014, 30 countries (14 from South America, 14 from Europe and 2 from Africa) entirely or partially privatised their public mandatory pension systems.

Taking into account the general principles of institutional money management, as well as the scrutiny and regulation related to the specificities of pension funds, in this research, the main goal is to test whether there are statistically significant differences in the risk-adjusted performance of the mandatory pension funds. The focus of this paper is on the former transition economies which participated in the reform of the pension system leading to the introduction of the fully-funded pension system pillar. Since the former transition economies in the sample have all been EU member states for more than ten years, they share not only broadly similar pension reform blueprints but also comparable structural features which enable easier comparison of the pension funds risk-categories performance. By comparing the performance of risk-categories A and B across countries, the research provides insight both regarding the existence of the differences within the same risk-category and between the two categories. The research findings are not only important from the regulatory policy perspective but are also related to other research efforts focusing on efficient portfolio diversification, i.e., advances in asset allocation for the institutional investors.

2 LITERATURE REVIEW

Analysis of pension funds' asset management can be placed in the context of financial research dealing with topics commonly related to institutional investors. Typically, therefore, a broad overview of research would include efficient portfolio diversification studies such as those by Amenc and Goltz (2013), Amenc et al. (2014), Cesarone, Scozzari and Tardella (2020) and Martellini and Milhau (2018), highlighting the need, on the one hand, for efficient exposure to different sources of systematic risk and, on the other, for effective elimination of idiosyncratic risk. In this context, the exact exposures to specific asset classes may vary substantially, while the ongoing ability to diversify efficiently is oriented towards maximising return per unit of risk in the portfolio. In this regard, Giron et al. (2018), in their analysis of retirement savings portfolio allocation, emphasize that bonds, relative to equities, will not provide the desired protection because managers may be unable to react to market changes. Anarkulova, Cederburg and O'Doherty (2025) go further, arguing that all retirement savings should be invested in equities, with 50% allocated to the domestic (U.S.) market and 50% to other markets, adding arguments in favour of the growing consensus on the so-called "death of the policy portfolio".

There are also specific topics related to retirement investing, focusing not just on the accumulation stage of the retirement investing problem but on providing the complete solution to the pension fund beneficiaries. The solution to the decumulation problem has been closely examined in the papers proposing retirement bonds, such as Kobor and Muralidhar (2018), and also in papers like Martellini, Milhau and Mulvey (2019) suggesting that smart target dates strategies can outperform what is currently being offered in the market. The recent performance of risk-categories A, B and C pension funds, for instance in Croatia, provides empirical evidence in support of this strand of research since the risk-category C pension funds were affected

PUBLIC SECTOR ECONOMICS
50 (3) 351-372 (2026)

by the inflation outburst in 2022 the most and were the slowest to recover in the following years. In other words, the manifestation of a rare combination of factors[2] resulted in risk-category A (considered the riskiest category) outperforming both B and C risk-categories during the fall in 2022 and during the following recovery years.[3] Thus, as pointed out by Martellini (2016) efficient diversification should be used in order to maximise the Sharpe ratio, i.e., maximise the return per unit of risk of the portfolio, but hedging is the appropriate risk management technique that should be used to protect the portfolio in times of crisis.

Other strands of research can primarily be classified into two groups. The first one, related to the countries that have introduced a fully-funded pension system pillar through pension reform, is concerned with whether the pension funds' management was able to deliver the expected risk-adjusted returns for the citizens. Such studies include research like the one performed by Mittelstaedt and Olsen (2003) for Chile, De la Torre Torres et al. (2015) for Mexico, Kompa and Witkowska (2015) for Poland, and Novaković (2015) for Croatia. The second one includes conducting cross-country studies encompassing research questions beyond the skill of a pension fund manager. Within this group, the work of Antolin (2008), Walker and Iglesias-Palau (2010) and Foo and Witkowska (2016) provide overviews of pension funds' performance in the US, Europe and OECD countries while discussing the role of regulation, investment strategies and even market structure when exploring the risk-adjusted performance differences across countries. Studies by Bohl, Lischewski and Voronkova (2011) and Kupčík and Gottwald (2017) also analyse the risk-adjusted performance in multiple countries, focusing on the assessment of differences and sources of outperformance. Also, Boon et al. (2014) and Boon et al. (2018) study a multicountry panel dataset and focus on the assessment of regulation on pension fund performance by employing panel regression models. Lastly, Afanador, Davis and Pedraza Morales (2021) focus on the effects of international diversification on performance, relying on the CAPM-based approach, showing that it increases returns, albeit with higher risk. Although there are economic reasons for "home bias," i.e., a preference for domestic investments, the study suggests that it is often driven by regulatory barriers and limited availability of instruments, which hinder international diversification. The paper, among other cases, analyses portfolios and returns in Bulgaria, Romania, Slovakia, and Chile, concluding that pension funds in Slovakia have the highest share of foreign investments, that this share is lower in Bulgaria and Romania, and that it is optimal in Chile.

All the above-mentioned research relies, among other risk-adjusted performance measures, on the Sharpe ratio. However, none of the research papers performing cross-country studies analyse the statistical significance of the Sharpe ratio differences by employing the testing methodology which is commonly used in research

[2] Prolonged period of low and even negative interest rates on the government bonds with the best credit rating followed by a surge in inflation.
[3] Available at: https://www.hanfa.hr/statistika/financijski-pokazatelji/mirex/?from=12.02.2026.&to=16.02.2026.

PUBLIC SECTOR ECONOMICS
50 (3) 351-372 (2026)

papers analysing the performance of investment funds, e.g., mutual funds and hedge funds. Papers, for instance, by Eling and Faust (2010) and Auer and Schuhmacher (2013) employing the Ledoit-Wolf and Jobson-Korkie-Memmel tests can be considered as such examples. Therefore, the contribution of this paper to the existing literature is twofold. On the one hand, this research follows Boon et al. (2014) in using the country-level aggregate pension funds' Sharpe ratios but employs the Jobson-Korkie-Memmel, Opdyke and Ledoit-Wolf tests in order to test the statistical significance of risk-adjusted performance of pension funds in the analysed countries. In this sense, this paper improves the analysis of pension funds' performance presented in the papers of, for instance, Bohl, Lischewski and Voronkova (2011) and Kupčík and Gottwald (2017). On the other hand, this paper improves the performance analysis not just by adding the statistical testing methodology but by conducting the analysis on a data sample which resembles the approach by Afanador, Davis and Pedraza Morales (2021).

The aim of the paper is, therefore, to compare the risk-adjusted performance of pension funds within the EU member states and particularly in the case of today's EU member states which belong to the group of former EU transition economies. Given that the sample in this study focuses on the transition economies (Slovakia, Romania, Bulgaria and Croatia), the potential findings related to differences in the pension funds' performance can in this case be more easily explained. Namely, considering the comparable legal frameworks, social protection arrangements, and the role and design of mandatory funded pension schemes, the potential differences for the same risk-categories of pension funds can be attributed mostly to the investment strategies employed or to the quality of management, which is an attempt to improve on the challenges discussed by Walker and Iglesias-Palau (2010) in their research. Moreover, in order to test whether there are performance differences between the pension funds of EU and non-EU countries, Chilean pension funds were included in the research sample. Also, the EU sample contains data on Swedish pension funds with Sweden being an EU member country that has not adopted the euro nor is it expected to adopt it in the foreseeable future. Thus, aside from Chilean funds, Swedish and Bulgarian[4] pension funds operate from outside the euro area, therefore making their Eurozone investments more international than in the case of the rest of the EU sample. Any detected performance differences between the pension funds from countries included in the sample can also be discussed in this context, following the research by Afanador, Davis and Pedraza Morales (2021). Also, in relation to the previous research, by comparing the risk-adjusted performance between different risk-categories of pension funds, this study provides additional evidence regarding the potential change in the regulatory framework as suggested by Anarkulova, Cederburg and O'Doherty (2025) and Giron et al. (2018).

[4] Bulgaria has joined the Eurozone on 1 January 2026, but the data sample includes data only up to December 2024.

3 DATA AND METHODS

3.1 DATA

As already explained in the previous section, the selection of countries included in the analysis was based on the objective of analysing pension systems that share comparable structural features and a common history of pension reform. According to assessments by international financial institutions and organisations, such as the World Bank (see, for instance, Hinz et al. (2010); OECD (2023), ILO (2018)) and additionally supported by "pension maps" presented in Schneider, Petrova and Becker (2021), Croatia, Slovakia, Romania, and Bulgaria exhibit substantial similarity in their legal frameworks, social protection arrangements and the role and design of mandatory funded pension schemes among EU transition economies. In addition, Sweden is included in the sample as a non-eurozone EU member state, which represents a mature, developed pension system that introduced funded pension savings several decades earlier than the former European transition economies. With one of the largest pension sectors in Europe relative to GDP, Sweden offers a relevant benchmark for comparison[5]. Chile is also included as a globally significant case of an extensive pension reform and a reference model for many European transition countries.

Pension funds in the selected countries are grouped into three risk-categories (A, B, and C) although they do not necessarily use those labels. For the purposes of this analysis, pension funds in the sample are classified either according to the target investment structure specified in each fund's prospectus or according to their current exposure to different asset classes, depending on which of these sources is available. Category A funds have a medium-to-high risk profile with high equity exposure (70 to 100%). Category B funds are moderately conservative, with risk between A and C. Category C funds are the most conservative, with little or no equity exposure (below 20%). This results in a sample consisting of 23 category A funds, 26 category B funds, and 18 category C funds. Given that, according to these classification criteria, there are no risk-category C funds in Romania and Bulgaria, the data and methodology described further, for the purpose of statistical risk-adjusted performance testing, is applied only to the risk-category A and B funds in pairwise tests.[6]

For each of these funds, monthly data on the unit value and the net asset value (NAV)[7] were collected for the period from July 2015 until 31 December 2024, and the corresponding time series of monthly returns and excess returns were calculated. Based on the time series of monthly excess returns, the Sharpe ratios[8] for each country's risk-category pension funds were computed in the pairwise testing

[5] Due to the unavailability of complete data for all Swedish pension funds, the analysis relies on a representative subset managed by Alecta, one of the largest pension providers in Europe.

[6] The available risk-category C data collected (at an annual frequency) was used for the purpose of initial analysis presented in table 1 at the beginning of the section of the paper presenting findings.

[7] Pension funds in Bulgaria and Sweden do not publish monthly NAV data. For these countries, OECD data on NAV, published on an annual basis, are used in the analysis.

[8] $SR = \frac{R_p - R_f}{\sigma_p}$, where R_p is the average return of the portfolio, R_f is the risk-free rate of return, and σ_p is the standard deviation of portfolio returns.

described in the following section. To enable meaningful cross-country comparisons, data on returns, i.e., excess returns are weighted by NAV, using weights based on each fund's NAV relative to the total NAV of all funds in the same category within that country. In this way, weighted averages are obtained that reflect the share of each fund's assets in the total assets of all funds belonging to the same category in a given country, with the notable exception in the case of Bulgaria and Sweden due to the fact that the NAV data for these two countries are only available at the annual frequency. Therefore, in the case of those two countries the monthly returns and excess returns calculation correspond to the calculations for other countries, but the weighted averages are calculated by equally weighting the monthly returns of funds in the risk-category. The risk-free rate is represented by the constantly available short-term reference rate on 3-month government securities.

3.2 METHODOLOGY

To test the differences in risk-adjusted performance, three tests designed specifically for the pairwise performance comparison of investments based on the Sharpe ratio are used. They enable the evaluation of whether the observed differences in risk-adjusted returns are statistically significant and allow for a systematic assessment of the relationship between the targeted risk profile of a fund and its realised outcomes, measured through returns and volatility. The same tests are used both to compare Sharpe ratios of the same category of pension funds across countries and between category A and category B funds within each country.

The first test employed is based on the following test statistic proposed by Jobson and Korkie (1981 and 1985):

$$z_{Sin} = \frac{\widehat{SR_{in}}}{\sqrt{\hat{\theta}}} \tag{1}$$

and θ is defined by the following equation:

$$\theta = \frac{1}{T}\left[2\sigma_i^2\sigma_n^2 - 2\sigma_i\sigma_n\sigma_{in} + \frac{1}{2}\mu_i^2\sigma_n^2 + \frac{1}{2}\mu_n^2\sigma_i^2 - \frac{\mu_i\mu_n}{2\sigma_i\sigma_n}\left(\sigma_{in}^2 + \sigma_i^2\sigma_n^2\right)\right] \tag{2}$$

where $\widehat{SR_{in}}$ is the Sharpe ratio difference[9] estimator between investments i and n, and θ is the variance of the asymptotic distribution of the Sharpe ratio statistic which is assumed to be normal and where μ denotes excess returns and σ standard deviation or correlation of the analysed investments and T is the sample size. Under the null hypothesis that the two Sharpe ratios are equal, a significant test value indicates a statistically meaningful difference in risk-adjusted performance between the investments, or in the case of this study – pension funds. In his paper, Memmel (2003) corrected the formula for θ which caused the formula to underestimate the asymptotic variance and simplified the formula[10], resulting in a formula

[9] In their paper, Jobson and Korkie (1981) actually refer to the transformed difference statistic, which is further explained in the original paper and is omitted here in order to be concise.

[10] $\rho_{i,j}$ in the formula denotes the correlation coefficient between the analysed investments.

for what is known as Jobson-Korkie-Memmel (JKM) test which is widely used in the finance literature for performance comparison. The JKM test formula as presented in, for instance, Auer and Schuhmacher (2013) is:

$$JKM = \frac{SR_i - SR_n}{\sqrt{\frac{1}{T}\left(2*\left(1-\rho_{ij}\right)+\frac{SR_i^2}{2}+\frac{SR_n^2}{2}-\rho_{ij}^{\,2}*SR_i*SR_n\right)}} \tag{3}$$

Due to the limitations of the JKM test, arising from the fact that the test is only valid for the normal distribution of returns and when returns do not exhibit serial correlation (as elaborated in detail in for example, Ledoit and Wolf, 2008), other tests were introduced. In this paper, therefore, the Opdyke (2007) test was also used which allows for non-normal return distributions and serially correlated returns. The test statistic is calculated by the standard formula referring to the Sharpe ratio difference estimator in the numerator and its standard error in the denominator:

$$Opdyke = \frac{\widehat{SR_i} - \widehat{SR_n}}{\sqrt{\frac{1}{T}*\left(\widehat{\sigma^2_{SR_i}}+\widehat{\sigma^2_{SR_n}}-2*Cov(SR_i,SR_n)\right)}} \tag{4}$$

where the symbols defined previously retain the same meaning. On the other hand, the $\widehat{\sigma^2_{SR}}$ estimators for the investments *i* and *n* are calculated according to a specific formula explicitly accounting for skewness and kurtosis in returns, therefore providing a generalization of earlier tests by relaxing the normality assumption which is presented and discussed in detail in Opdyke (2007). Finally, the Ledoit-Wolf test assesses the statistical significance of the differences between the two investments by addressing the limitations of both non-normality and serial correlation of returns, which is appropriate even in the case of smaller data samples. The test is based on the same standard formula for the test statistic as presented in equation (4). However, unlike in the previous tests, in order to provide a robust standard error, the standard error estimator is not provided analytically but is calculated by relying on the time-series block bootstrap approach as presented in Ledoit and Wolf (2008): 853-855[11].

4 RESEARCH FINDINGS

This section of the paper is divided into two parts. The first part performs initial analysis and reports descriptive statistics and tests related to the normality of distribution and serial correlation of the analysed data. The second part presents the main research findings related to the testing of the Sharpe ratio differences.

[11] In the same paper, Ledoit and Wolf (2008) also propose and test the Heteroskedasticity and Autocorrelation Consistent (HAC) which is similar to the Opdyke (2007) test. Due to the mentioned similarity and conclusions in the Ledoit and Wolf (2008) paper, which favour the time-series block bootstrap approach for the typical (smaller) financial sample sizes, in this paper we refer to the Ledoit and Wolf (LW) test and use the time-series block bootstrap approach for calculating the standard error.

4.1 INITIAL ANALYSIS

To provide an initial overview of pension fund performance, table 1 presents returns, volatilities, and Sharpe ratios for the analysed period across the three fund categories. This descriptive comparison serves as the basis for the subsequent significance testing of Sharpe ratio differences.

Table 1

Annual average return, volatility, and Sharpe ratio of pension funds for the analysed period

	Categories								
	A			B			C		
Country	Return (%)	Volatility (%)	Sharpe ratio	Return (%)	Volatility (%)	Sharpe ratio	Return (%)	Volatility (%)	Sharpe ratio
HR	8.31	5.92	1.12	4.85	3.76	1.09	3.55	2.26	1.07
SVK	9.46	16.24	0.60	4.37	8.97	0.44	0.34	2.33	-0.21
RO	5.67	5.36	0.50	5.34	5.21	0.45	–	–	–
BG	2.82	3.76	0.68	2.12	4.96	0.37	–	–	–
CHI	7.98	11.67	0.59	7.24	8.12	0.72	6.07	7.80	0.60
SWE	5.68	8.61	0.77	4.66	6.72	0.72	2.01	3.13	0.50

Note: Country abbreviations are as follows: HR – Croatia, SVK – Slovakia, RO – Romania, BG – Bulgaria, CHI – Chile, SWE – Sweden.

Source: Authors' research.

Based on the results in table 1, it can be seen that returns are always the highest in risk-category A and the lowest in risk-category C, while volatility is also the highest in risk-category A and the lowest in risk-category C, except in the case of Bulgarian pension funds, where volatility is higher in risk-category B than in risk-category A. Sharpe ratios are higher across the analysed countries in risk-category A compared with risk-categories B and C, and are the lowest in risk-category C with the exception of Chilean funds, where risk-category B funds achieved a higher Sharpe ratio than risk-category A funds. Also, the Sharpe ratios for all three risk-categories seem to be substantially higher in the case of Croatian pension funds relative to other countries. In risk-category A Swedish and Bulgarian funds also stand out, while Swedish and Chilean funds stand out in the risk-category B. Risk-category C, unfortunately, lacks data for the Romanian and Bulgarian funds as already mentioned in section 2.1. Moreover, as shown here, the Slovakian risk-category C funds stand out with a negative Sharpe ratio.

However, in order to draw a conclusion based on the collected dataset, the next step of the research assesses whether the differences identified in the initial analysis are statistically significant by conducting the tests presented in the section of the paper referring to the methodology (2.2). Considering the importance of the data characteristics regarding the proposed testing, the findings regarding normality and serial correlation of analysed time series are presented in tables 2 and 3. The results of the conducted Shapiro-Wilk and Jarque-Bera tests in table 2 both clearly show that only two (A and B risk-category pension funds in Sweden,

denoted by SWE A and SWE B) out of the twelve-time series of excess returns in the data sample seem to be normally distributed.

Table 2

Results of normality testing based on Shapiro-Wilk and Jarque-Bera tests

Country and risk-category	Skewness	Excess Kurtosis	Shapiro-Wilk Stat	SW p-value	Jarque-Bera Stat	JB p-value
HR A	-1.858	7.989	0.880***	(0.000)	365.596***	(0.000)
HR B	-1.260	2.999	0.923***	(0.000)	72.252***	(0.000)
SVK A	-1.135	6.449	0.892***	(0.000)	220.078***	(0.000)
SVK B	-1.677	10.303	0.840***	(0.000)	552.806***	(0.000)
RO A	-0.543	2.316	0.962***	(0.003)	30.808***	(0.000)
RO B	-0.567	2.183	0.963***	(0.003)	28.498***	(0.000)
BG A	0.048	3.562	0.935***	(0.000)	59.784***	(0.000)
BG B	-1.464	4.924	0.898***	(0.000)	154.514***	(0.000)
CHI A	-0.714	4.222	0.938***	(0.000)	93.511***	(0.000)
CHI B	-0.692	4.255	0.947***	(0.000)	94.247***	(0.000)
SWE A	-0.313	0.347	0.987	**(0.369)**	2.415	**(0.299)**
SWE B	-0.336	0.386	0.984	**(0.195)**	2.834	**(0.242)**

*Note: Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*
Source: Authors' research.

Regarding serial correlation, the analysed time series exhibit better performance with the first-order serial correlation being present only in the A and B risk-category pension funds in Slovakia (at 1% and 5% statistical significance levels, respectively, denoted by SVK A and SVK B) and in the case of Chilean risk-category A pension funds (at 10% statistical significance, denoted by CHI A).

Table 3

Results of serial correlation testing

Country and risk-category	Durbin-Watson	ACF (Lag 1)	Ljung-Box Stat	LB p-value
HR A	1.835	-0.030	0.103	(0.748)
HR B	1.849	-0.029	0.099	(0.752)
SVK A	2.483	-0.294	10.040***	(0.002)
SVK B	2.428	-0.235	6.412**	(0.011)
RO A	1.863	0.046	0.248	(0.618)
RO B	1.866	0.049	0.273	(0.601)
BG A	2.026	-0.104	1.260	(0.262)
BG B	1.856	0.051	0.305	(0.581)
CHI A	2.266	-0.168	3.284*	(0.070)
CHI B	2.138	-0.118	1.626	(0.202)
SWE A	2.055	-0.085	0.839	(0.360)
SWE B	2.075	-0.093	0.997	(0.318)

*Note: Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*
Source: Authors' research.

4.2 SHARPE RATIO PERFORMANCE TESTING

The first part of the Sharpe ratio differences testing is related to the performance testing between the same risk-categories of funds across the analysed countries. The results of the pairwise testing are therefore presented in matrix form[12] in tables 4 and 5 – with table 4 referring to the results of tests for risk-category A and table 5 for risk-category B.

Table 4

Results of Sharpe ratio differences testing for risk-category A

	HR	SVK	RO	BG	CHI	SWE
HR	–	JKM: 1.976** (0.048) Opdyke: 1.925* (0.054) LW: 3.187** (0.010)	JKM: 2.505** (0.012) Opdyke: 2.069** (0.039) LW: 3.360** (0.012)	JKM: 1.211 (0.226) Opdyke: 1.197 (0.231) LW: 1.667 (0.108)	JKM: 1.995** (0.046) Opdyke: 1.947* (0.052) LW: 2.010* (0.060)	JKM: 1.964** (0.050) Opdyke: 1.483 (0.138) LW: 1.775 (0.104)
SVK		–	JKM: 0.337 (0.736) Opdyke: 0.392 (0.695) LW: 0.466 (0.658)	JKM: -0.478 (0.633) Opdyke: -0.668 (0.504) LW: -0.747 (0.494)	JKM: 0.149 (0.881) Opdyke: 0.219 (0.827) LW: 0.271 (0.812)	JKM: -0.294 (0.769) Opdyke: -0.399 (0.690) LW: -0.636 (0.562)
RO			–	JKM: -0.922 (0.356) Opdyke: -1.106 (0.269) LW: -1.618 (0.114)	JKM: -0.210 (0.834) Opdyke: -0.208 (0.835) LW: -0.257 (0.782)	JKM: -0.707 (0.479) Opdyke: -0.708 (0.479) LW: -0.730 (0.544)
BG				–	JKM: 0.596 (0.551) Opdyke: 0.754 (0.451) LW: 0.862 (0.424)	JKM: 0.227 (0.820) Opdyke: 0.276 (0.782) LW: 0.321 (0.750)
CHI					–	JKM: -0.500 (0.617) Opdyke: -0.582 (0.560) LW: -0.951 (0.466)
SWE						–

Note: Each cell reports the test statistic followed by the corresponding p-value in brackets (with JKM referring to the Jobson-Korkie-Memmel and LW to the Ledoit-Wolf test).

*Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

Source: Authors' research.

[12] Due to symmetry, only half of the results in the table are reported.

The test results in table 4 show that a statistically significant difference in the Sharpe ratios in risk-category A is predominantly present in the case of Croatian pension funds (HR). In the analysed time period, the difference seems to be statistically significant relative to the Slovak (SVK), Romanian (RO) and Chilean (CHI) pension funds based on all three tests and at least at the 10% level. In the case of the relative performance of Croatian and Swedish pension funds both Opdyke and Ledoit-Wolf tests do not find the Sharpe ratio statistically significantly different at the usual significance levels, which is important due to the non-normal distribution of the data related to the Croatian pension funds. All the other pairwise tests, surprisingly, do not provide any evidence of statistically significant performance differences when measured by the Sharpe ratio in the analysed time period.

Table 5

Results of Sharpe ratio differences testing for risk-category B

	HR	SVK	RO	BG	CHI	SWE
HR	–	JKM: 2.370** (0.018) Opdyke: 2.725*** (0.006) LW: 2.044* (0.088)	JKM: 2.279** (0.023) Opdyke: 2.106** (0.035) LW: 2.280** (0.032)	JKM: 3.160*** (0.002) Opdyke: 2.943*** (0.003) LW: 2.843** (0.012)	JKM: 1.226 (0.220) Opdyke: 1.141 (0.254) LW: 1.228 (0.274)	JKM: 1.766* (0.077) Opdyke: 1.532 (0.125) LW: 1.640* (0.090)
SVK		–	JKM: -0.052 (0.959) Opdyke: -0.063 (0.950) LW: -0.092 (0.922)	JKM: 0.108 (0.914) Opdyke: 0.140 (0.889) LW: 0.160 (0.916)	JKM: -0.920 (0.358) Opdyke: -1.160 (0.246) LW: -1.268 (0.270)	JKM: -0.813 (0.416) Opdyke: -1.011 (0.312) LW: -1.375 (0.350)
RO			–	JKM: 0.169 (0.866) Opdyke: 0.167 (0.867) LW: 0.239 (0.816)	JKM: -0.827 (0.408) Opdyke: -0.781 (0.435) LW: -0.849 (0.432)	JKM: -0.803 (0.422) Opdyke: -0.812 (0.417) LW: -0.816 (0.520)
BG				–	JKM: -1.114 (0.265) Opdyke: -1.173 (0.241) LW: -0.974 (0.390)	JKM: -1.379 (0.168) Opdyke: -1.363 (0.173) LW: -1.240 (0.290)
CHI					–	JKM: 0.101 (0.919) Opdyke: 0.107 (0.915) LW: 0.107 (0.908)
SWE						–

Note: Each cell reports the test statistic followed by the corresponding p-value in brackets (with JKM referring to the Jobson-Korkie-Memmel and LW to the Ledoit-Wolf test).

*Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

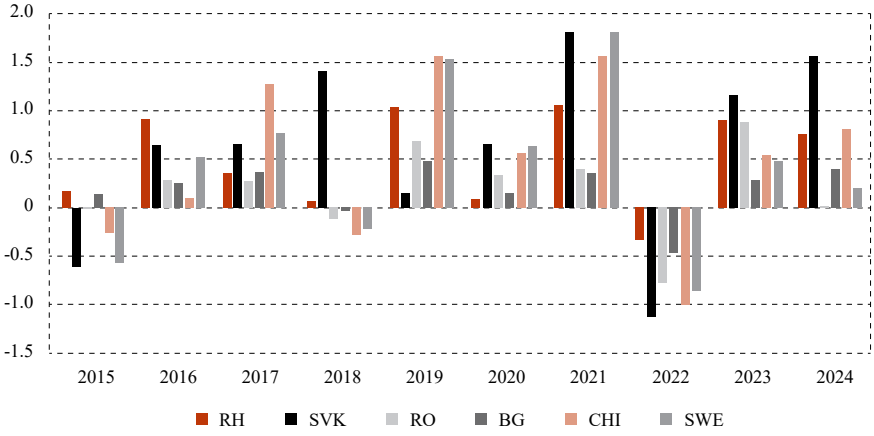
Source: Authors' research.

As shown in table 5, the test results for risk-category B funds are quite similar to those in the previous analysis, with statistically significant differences in Sharpe ratios found in all instances relative to Croatian pension funds (HR). The difference in this case is that no significant difference is detected relative to Chilean funds (CHI), but a significant difference relative to Bulgarian funds (BG) emerges across all three tests (most notably on Ledoit-Wolf at the 5% significance level). Also, even the difference relative to Swedish funds (SWE) proved to be statistically significant at the 10% level (albeit just barely with the p-value of 0.09).

In order to further explore the source of outperformance among Croatian pension funds, based solely on the available dataset of excess returns time series, the following figures 1 and 2 present the breakdown of pension funds' performance per country and year for each analysed risk-category. As can be seen in both charts, Croatian pension funds outperform other countries in the sample on average in 2015, 2016 and 2022[13] in terms of excess return. In other years, recorded returns are better than those of most other countries, leading to lower overall volatility over the analysed period and resulting in outperformance.

Figure 1

Average excess returns for risk-category A funds by country and year (in %)



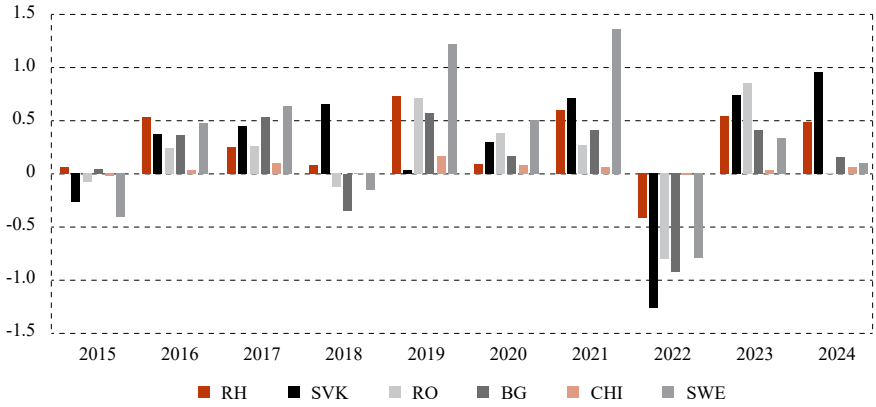
Source: Authors' research.

[13] In 2022 the highest excess return is actually related to the lowest loss.

PUBLIC SECTOR ECONOMICS
50 (3) 351-372 (2026)

FIGURE 2

Average excess returns for risk-category B funds by country and year (in %)



Source: Authors' research.

The performance presented in figures 1 and 2 can be briefly supplemented by elaborating on the evolution of the structure of investments of Croatian pension funds. When looking at the structure of assets of Croatian pension funds in table 6, it can be noted that at the end of 2015 and 2016 the structure of assets of both risk-categories of Croatian pension funds was vastly dominated by domestic investments. For risk-category A, they were around 80%, while for risk-category B they were well above the 80%. The difference between the two was in the composition, with domestic government bonds accounting for more than 70% of assets in the case of risk-category B, while holding at close to but below 50% in the case of risk-category A. The rest was primarily invested in the domestic stock market, which therefore accounted for just slightly more than 10% in risk-category B, but on the other hand made up 20-30% for risk-category A.

TABLE 6

The structure of assets of Croatian pension funds (% of total assets, end of period)

	Risk-category							
	A				B			
Investments*	**2015**	**2016**	**2022**	**2024**	**2015**	**2016**	**2022**	**2024**
Domestic	**78**	**81**	**61**	**51**	**85**	**88**	**73**	**62**
Shares and GDRs	23	31	24	17	11	13	13	15
Government bonds	49	46	29	24	71	71	57	43
Foreign	**21**	**14**	**30**	**47**	**14**	**10**	**24**	**37**
Shares	12	12	12	14	9	7	9	9
Government bonds	0	0	4	13	0	0	4	15
Undertakings for collective investment in transferable securities and open investment funds	8	2	12	14	5	2	9	9

**Both categories include more asset classes than shown in this table. They were omitted in order to present the changes in the dominant asset classes more clearly.*

Source: Authors' research based on data from the Croatian Financial Services Supervisory Agency.[14]

[14] Data available for download at: https://www.hanfa.hr/statistika/mirovinski-fondovi/.

For both risk-categories, the share of domestic investments decreased steadily in the following years dropping to slightly above 60% and 50% at the end of 2022 and 2024, respectively, for risk-category A and close to 70% and 60% for the same years for risk-category B. However, in the same period exposure to domestic stocks increased steadily in risk-category B reaching 15% at the end of 2024 while decreasing steadily to 17% at the end of the same year for risk-category A. At the same time, exposure to foreign markets in terms of foreign stock investments remained roughly stable, in the 12-14% range for risk-category A and 7-9% for risk-category B. However, both risk-categories increased exposure to foreign government bonds from 0% during the 2015-2016 period to 4% at the end of 2022 and eventually to 13% for the risk-category A and 15% for the risk-category B at the end of 2024. Furthermore, the share of foreign investments held through open investment funds rose from 8% to 14% for risk-category A and from 5% to 9% for risk-category B between the end of 2015 and the end of 2024.

Based on the provided data, the most likely factor driving the performance of Croatian pension funds is the exposure to domestic government bonds. This is supported by the fact that Croatia returned to the investment grade credit rating in 2019 which was followed by the COVID-19 pandemic in 2020. It is less clear how much direct investment in foreign government bonds and exposure to foreign markets through open investment funds contributed to performance in the post-pandemic period.

The second part of the Sharpe ratio differences testing is related to the performance testing between A and B fund risk-categories within the same country. The results of the pairwise testing for the six analysed countries are presented in table 7 below.

Table 7

Results of Sharpe ratio differences testing within countries – risk-category A vs. B

Country	JKM	Opdyke	LW	Sharpe ratio difference
HR A vs. HR B	0.611 (0.542)	0.427 (0.670)	0.428 (0.708)	0.067
SVK A vs. SVK B	1.486 (0.137)	1.396 (0.163)	1.349 (0.442)	0.177
RO A vs. RO B	1.465 (0.143)	1.382 (0.167)	1.463 (0.136)	0.047
BG A vs. BG B	1.363 (0.173)	1.773* (0.076)	2.175** (0.032)	0.374
CHI A vs. CHI B	-1.053 (0.292)	-1.013 (0.311)	-1.087 (0.268)	-0.146
SWE A vs. SWE B	0.548 (0.584)	0.534 (0.593)	0.477 (0.674)	0.019

Note: Each cell reports the test statistic followed by the corresponding p-value brackets (with JKM referring to the Jobson-Korkie-Memmel and LW to the Ledoit-Wolf test).

*Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

Source: Authors' research.

PUBLIC SECTOR ECONOMICS
50 (3) 351-372 (2026)

The reported results show that, in the analysed period, the p-values are greater than the 0.1 threshold for all the analysed countries except for Bulgaria (BG A vs BG B), for which a statistically significant difference is detected based on the Opdyke and Ledoit-Wolf tests at the 10% and 5%, respectively. For other countries, even in the case of Slovakia (SVK), which exhibits the second largest difference in Sharpe ratios of 0.1773 (after the 0.3736 for Bulgaria), no evidence of a statistically significant difference between the performance of risk-category A vs. risk-category B funds, as measured by the Sharpe ratio, is found. Thus, the differences in the target risk profile between these two categories do not provide evidence of a statistically significant difference in risk-adjusted returns for five out of six analysed countries.

In order to provide a comprehensive analysis, table 8 reports Sharpe ratio differences for all the conducted pairwise tests, along with corresponding significance level of the difference according to the Ledoit-Wolf test. The table clearly demonstrates the role that the Sharpe ratio difference plays in the statistical significance testing. In this sense the magnitude of outperformance of the Croatian pension funds relative to other countries as the underlying factor of statistical significance of Sharpe ratio differences becomes obvious. Results in table 8 reveal that, therefore, as expected, the Sharpe ratios of Croatian pension funds of risk-category A significantly differ from those of other countries' risk-category B funds due to, generally, an even larger difference than in the case when the same risk-categories were compared. When the same risk-categories were compared, the Sharpe ratio differences stood in the range of 0.5568 – 0.6713 and 0.4174 – 0.6929 for the risk-categories A and B, respectively. When Croatian pension fund risk-category A and other countries' risk-category B are compared, the difference spikes above the 0.7 level with the exceptions of Sweden, for which the difference stands at 0.4845, and for Chile,[15] with a difference of 0.4565, which is not even found to be statistically significant. Lastly, it can be noted that even the Sharpe ratio differences between the Croatian pension funds' risk-category B show large and, therefore, statistically significant differences when compared to the Slovakian and Romanian risk-category A funds. It is interesting to note that even when risk-category A is compared to risk-category B there are no recorded cases of a statistically significant difference in any of the analysed pairwise tests across countries in the sample with Sharpe ratio differences rarely reaching the magnitude of 0.3 in absolute terms.

[15] As can be noted in table 1, the Sharpe ratio of risk-category B in Chile is higher than that of risk-category A. This anomaly causes the difference in Sharpe ratios between Croatian and Chilean pension funds of risk-category A to be statistically significant, while proving insignificant when risk-categories A and B of the same countries are compared.

Table 8

Sharpe ratio differences and their statistical significance (based on the Ledoit-Wolf test)

Country and risk-category	HR A	HR B	SVK A	SVK B	RO A	RO B	BG A	BG B	CHI A	CHI B	SWE A	SWE B
HR A	–	0.067	**0.557****	**0.734****	**0.671****	**0.718****	0.386	**0.760*****	**0.603***	0.457	0.466	**0.485***
HR B		–	**0.490***	**0.667***	**0.604***	**0.651****	0.319	**0.693****	0.536	0.389	0.399	**0.417***
SVK A			–	0.177	0.115	0.161	-0.170	0.203	0.046	-0.100	-0.091	-0.072
SVK B				–	-0.063	-0.016	-0.348	0.026	-0.131	-0.278	-0.268	-0.250
RO A					–	0.047	-0.285	0.089	-0.068	-0.215	-0.205	-0.187
RO B						–	-0.331	0.042	-0.115	-0.262	-0.252	-0.234
BG A							–	**0.374****	0.217	0.070	0.080	0.098
BG B								–	-0.157	-0.304	-0.294	-0.276
CHI A									–	-0.146	-0.137	-0.118
CHI B										–	0.010	0.028
SWE A											–	0.019
SWE B												–

Note: Each cell reports the Sharpe ratio differences for row vs column. The lower half of the matrix would contain the same values with the opposite signs if shown.

*Asterisks ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively, referring to the Ledoit-Wolf test.*

No test statistic or the corresponding p-value is shown in this table as they are presented in tables 4 through 6 earlier in the text.

Source: Authors' research.

PUBLIC SECTOR ECONOMICS
50 (3) 351–372 (2026)

5 CONCLUSION

The research study presented in this paper focuses on the performance of pension funds by comparing their risk-adjusted performance based on the Sharpe ratio at the country level for the A and B pension fund risk-categories. Studies in this field are not only important due to the frequent criticisms of the pension reform in the transition economies and the general concerns related to the sustainability of fully funded (nowadays primarily defined contribution) pension systems, but also in an attempt to shed light on possible differences in the pension fund management across countries and in order to attempt to steer the improvement of the industry's regulatory framework. In this sense, the research tested Sharpe ratio differences by comparing the performance of the same risk-categories of pension funds across countries and within each country, by comparing performances between the two risk-categories. The research findings for the six analysed countries point out the statistically significant Sharpe ratio differences in the case of Croatian pension funds relative to other countries in both analysed risk-categories. The detected statistically significant outperformance in the analysed period stems from relatively high average excess returns which exhibit overall low volatility in the analysed period relative to the other countries. In the case of other countries, no statistically significant difference in performance is found, even when the risk-category A versus B funds across other countries are compared. When comparing risk-categories A and B within each country, the statistically significant difference in Sharpe ratios is found only in the case of Bulgaria highlighting the lack of evidence that would support the existence of a difference in risk-adjusted returns between the A and B risk-categories, regardless of their different risk profile target.

The findings of this study point to the surprising finding that there seems to be no major evidence of differences in risk-adjusted performance across countries, even when comparing the five EU member states in the sample with Chile or the four former transition economies in Eastern Europe in the sample with Sweden. It would, however, be worthwhile looking into the evolution of the portfolio structure of the Croatian pension funds and other analysed countries in order to better understand the sources of the detected statistically significant differences between Croatian pension funds and other countries. Such analysis would provide additional evidence that could be used to draw further conclusions about factors driving the pension funds' performance. Additionally, it could be noted that higher returns by pension funds in Slovakia and Chile, which are accompanied by higher volatilities, as shown in the initial analysis in this study (table 1), could be explained by a greater share of foreign investments (international diversification) based on the findings by Afanador, Davis and Pedraza Morales (2021). However, as shown in the initial analysis, the increased returns are offset by an increase in volatility and result in similar Sharpe ratios, for which consequently no significant differences are found in testing. Therefore, no effect noted by Afanador, Davis and Pedraza Morales (2021) can be corroborated in this research. Furthermore, the consistent lack of statistically significant differences in performance between A and B risk-categories within countries (with the notable exception of Bulgaria)

implies that future changes to the regulatory framework could move away from the current understanding of the system of risk-categories, corroborating the views in Giron et al. (2018) and Anarkulova, Cederburg and O'Doherty (2025) who advocate abandoning the "policy portfolio". In this sense, risk-categories focusing on providing efficient diversification should strive to achieve the highest possible Sharpe ratio. While it could be argued that it is possible to achieve Sharpe ratios of similar values related to different risk-categories such that one risk-category exhibits higher volatility (and higher corresponding return) and the other exhibits lower volatility (and lower corresponding return), even low-risk efficiently diversified portfolios fail to offer protection in times of crisis because hedging (not efficient diversification) is the appropriate risk management technique to address the issue as pointed out by Martellini (2016). Although the empirical analysis in this research did not include the risk-category C funds due to the lack of data, their recent performance in Croatia, as discussed in the literature review, provides empirical evidence consistent with this observation.

The limitations of this research include primarily the relatively small sample related to the number of countries included, but also the monthly data available related to the period of slightly less than ten years. Future research could expand both the monthly time series and the number of countries by including more developed countries in the sample and by making the sample more geographically diverse. Furthermore, when undertaking this task, future research should take into account the limitations of relying on the Sharpe ratio for such analysis, as elaborated in Walker and Iglesias-Palau (2010), given that the levels of risk-free interest rates and investment opportunities differ across large datasets. Also, future research could, for this purpose, rely on panel regressions instead of pairwise tests, as in Boon et al. (2014) and Boon, Brière and Rigot (2018). The results of such future research, together with the future research focused on the analysis of portfolio structure across different countries, could provide a comprehensive overview of factors driving the performance of pension funds. This would help design a more efficient and sustainable fully-funded pension system, increasing public confidence in the overall pension system.

Disclosure statement

The authors have no conflicts of interest to declare.

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

No generative AI or AI-assisted tools were used in the preparation of this manuscript.

REFERENCES

1. Afanador, J. P., Davis, R. M. and Pedraza Morales, A. E., 2021. Estimating the Gains from International Diversification: The Case of Pension Funds. *Policy Research Working Paper Series,* No. 9635. https://doi.org/10.1596/1813-9450-9635
2. Amenc, N. and Goltz, F., 2013. Smart Beta 2.0. *The Journal of Index Investing,* 4(3), pp. 15-23. https://doi.org/10.3905/JII.2013.4.3.015
3. Amenc, N. [et al.], 2014. Towards Smart Equity Factor Indices: Harvesting Risk Premia Without Taking Unrewarded Risks. *The Journal of Portfolio Management,* 40(4), pp. 106-122. https://doi.org/10.3905/jpm.2014.40.4.106
4. Anarkulova, A., Cederburg, S. and O'Doherty, M. S., 2025. *Beyond The Status Quo: A Critical Assessment of Lifecycle Investment Advice*. http://dx.doi.org/10.2139/ssrn.4590406
5. Andonov, A., Bauer, R. and Cremers, M., 2012. *Can Large Pension Funds Beat the Market? Asset Allocation, Market Timing, Security Selection and the Limits of Liquidity.* http://dx.doi.org/10.2139/ssrn.1885536
6. Antolin, P., 2008. Pension Fund Performance. *OECD Working Papers on Insurance and Private Pensions*, No. 20. https://doi.org/10.1787/240401404057
7. Auer, B. R. and Schuhmacher, F., 2013. Performance Hypothesis Testing With the Sharpe Ratio: The Case of Hedge Funds. *Finance Research Letters,* 10(4), pp. 196-208. https://doi.org/10.1016/j.frl.2013.08.001
8. Blake, D., Lehmann, B. N. and Timmermann, A., 1999. Asset Allocation Dynamics and Pension Fund Performance. *The Journal of Business,* 72(4), pp. 429-461.
9. Bohl, M. T., Lischewski, J. and Voronkova, S., 2011. Pension Funds' Performance in Strongly Regulated Industries in Central Europe: Evidence from Poland and Hungary. *Emerging Markets Finance and Trade,* 47(3), pp. 80-94. https://doi.org/10.2753/REE1540-496X4704S305
10. Boon, L.-N. [et al.], 2014. *Pension Regulation and Investment Performance: Rule-Based Vs. Risk-Based*. http://dx.doi.org/10.2139/ssrn.2400534
11. Boon, L. N., Brière, M. and Rigot, S., 2018. Regulation and Pension Fund Risk-Taking. *Journal of International Money and Finance,* 84, pp. 23-41. https://doi.org/10.1016/j.jimonfin.2018.01.005
12. Brinson, G. P., Singer, B. D. and Beebower, G. L., 1991. Determinants of Portfolio Performance II: An Update. *Financial Analysts Journal,* 47(3), pp. 40-48. https://doi.org/10.2469/faj.v47.n3.40
13. Broeders, D. and De Haan, L., 2020. Benchmark Selection and Performance. *Journal of Pension Economics and Finance,* 19(4), pp. 511-531. https://doi.org/10.1017/S1474747219000246
14. Cesarone, F., Scozzari, A. and Tardella, F., 2020. An Optimization-Diversification Approach to Portfolio Selection. *Journal of Global Optimization,* 76, pp. 245-265. https://doi.org/10.1007/s10898-019-00809-7
15. De la Torre Torres, O. V. [et al.], 2015. A Minimum Variance Benchmark to Measure the Performance of Pension Funds in Mexico. *Contaduría y Administración,* 60(3), pp. 593-614. https://doi.org/10.1016/j.cya.2015.05.009

16. Dyck, A., Manoel, P. and Morse, A., 2022. Outraged by Compensation: Implications for Public Pension Performance. *The Review of Financial Studies,* 35(6), pp. 2928-2980. https://doi.org/10.1093/rfs/hhab109
17. Eling, M. and Faust, R., 2010. The Performance of Hedge Funds and Mutual Funds in Emerging Markets. *Journal of Banking & Finance,* 34(8), pp. 1993-2009. https://doi.org/10.1016/j.jbankfin.2010.01.008
18. Foo, J. and Witkowska, D., 2016. Investment Performance of US and European Pension Funds. A Comparative Analysis. *The IEB International Journal of Finance,* 13, pp. 99-109.
19. Giron, K. [et al.], 2018. *Applying Goal-Based Investing Principles to The Retirement Problem*. Singapore: EDHEC-Risk Institute.
20. Hinz, R. [et al.], 2010. *Evaluating the Financial Performance of Pension Funds*. Washington, DC: The World Bank.
21. ILO, 2018. *The ILO Multi-Pillar pension model: Building equitable and sustainable pension systems*. Geneva: International Labour Organization.
22. Jobson, J. D. and Korkie, B., 1985. Some tests of linear asset pricing with multivariate normality. *Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences De L'administration,* 2(1), pp. 114-138. https://doi.org/10.1111/j.1936-4490.1985.tb00397.x
23. Jobson, J. D. and Korkie, B., 1981. Performance Hypothesis Testing with the Sharpe and Treynor Measures. *The Journal of Finance,* 36(4). pp. 889-908. https://doi.org/10.1111/j.1540-6261.1981.tb04891.x
24. Kobor, A. and Muralidhar, A., 2018. *How a New Bond Can Greatly Improve Retirement Security*. http://dx.doi.org/10.2139/ssrn.3148747
25. Kompa, K. and Witkowska, D., 2015. Pension System in Poland: Performance of Pension Funds. *Estudios De Economía Aplicada,* 33(3), pp. 965-984. https://doi.org/10.25115/eea.v33i3.3142
26. Kupčík, P. and Gottwald, P., 2017. The Return-Risk Performance of Selected Pension Fund in OECD With Focus on the Czech Pension System. *Acta Universitatis Agriculturae Et Silviculturae Mendelianae Brunensis,* 64(6), pp. 1981-1988. https://doi.org/10.11118/actaun201664061981
27. Ledoit, O. and Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance,* 15, pp. 850-859. https://doi.org/10.1016/j.jempfin.2008.03.002
28. Martellini, L., 2016. Mass customization versus mass production – how an industrial revolution is about to take place in money management and why it involves a shift from investment products to investment solutions. *Journal of Investment Management,* 14(3), pp. 5-13.
29. Martellini, L. and Milhau, V., 2018. Proverbial Baskets Are Uncorrelated Risk Factors! A Factor-Based Framework For Measuring And Managing Diversification In Multi-Asset Investment Solutions. *Journal of Portfolio Management, Suppl. Multi-Asset Special Issue,* 44(2), pp. 8-22. https://doi.org/10.3905/jpm.2018.44.2.008

30. Martellini, L., Milhau, V. and Mulvey, J., 2019. "Flexicure" retirement solutions: A part of the answer to the pension crisis? *Journal of Portfolio Management,* 45(5), pp. 136-151. https://doi.org/10.3905/jpm.2019.45.5.136
31. Memmel, C., 2003. Performance hypothesis testing with the Sharpe ratio. *Finance Letters,* (1), pp. 21-23.
32. Mittelstaedt, H. F. and Olsen, J. C., 2003. An empirical analysis of the investment performance of the Chilean pension system. *Journal of Pension Economics and Finance,* 2(1), pp. 7-24. https://doi.org/10.1017/S1474747202001208
33. Novaković, D., 2015. Evaluation Of The Financial Performance Of Pension Funds In Croatia. *Ekonomski vjesnik / Econviews – Review of Contemporary Business, Entrepreneurship and Economic Issues,* 28(1), pp. 199-212.
34. OECD, 2023. *Pensions at a Glance 2023*. Paris: OECD. https://doi.org/10.1787/678055dd-en
35. Opdyke, J. D., 2007. Comparing Sharpe ratios: So where are the *p*-values? *Journal of Asset Management,* 8, pp. 308-336. https://doi.org/10.1057/palgrave.jam.2250084
36. Schneider, S., Petrova, T. and Becker, U., 2021. *Pension Maps: Visualising the Institutional Structure of Old Age Security in Europe.* Munich: Mpisoc. https://doi.org/10.17617/2.3359088
37. Walker, E. and Iglesias-Palau, A., 2010. *Financial Performance of Pension Funds: An Exploratory Study*. https://doi.org/10.1596/978-0-8213-8159-5

Household behavioural response to health policy recommendations: evidence from the COVID-19 pandemic in Bosnia and Herzegovina, Croatia, Serbia and Slovenia

ROBERT J. SONORA, Ph.D.*
JOSIP TICA, Ph.D.*

Article**
JEL: C22, E70, H75, I12, I18
https://doi.org/10.3326/pse.50.3.3

* We would like to thank the participants of the weekly Brown Bag seminar at the Faculty of Economics at the University of Zagreb, as well as attendees of the seminar at the Institute of Public Finance, for their thoughtful comments and helpful feedback. We also appreciate comments from the editor and two anonymous referees which undoubtedly improved the quality of the paper. Finally, we are grateful to Katarina Ott for her valuable comments and encouragement in completing this paper.

** Received: March 31, 2025
Accepted: June 15, 2026

Robert J. SONORA
Accounting and Finance Department, University of Montana, 32 Campus Drive, Missoula, MT, 58912, USA
e-mail: robert.sonora@mso.umt.edu
ORCiD: 0000-0003-4713-7082

Josip TICA
Faculty of Economics and Business, University of Zagreb, Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia
e-mail: jtica@efzg.hr
ORCiD: 0000-0001-7937-1573

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

Abstract

The COVID-19 pandemic created an environment where rapidly changing policy interventions and heightened uncertainty substantially influenced individual behavioural responses. This study investigates the relationship between government mobility restrictions and public adherence in Bosnia and Herzegovina, Croatia, Serbia and Slovenia using mobility data and policy stringency indices. Focusing on the time-series properties of behavioural adjustment, we evaluate whether individual responses converge towards policy objectives over time. In Croatia, Bosnia and Herzegovina, and Slovenia, individuals initially exhibited stronger mobility reductions than implied by policy measures, consistent with temporary behavioural overreaction and heightened short-run risk sensitivity. In Serbia, behavioural responses initially remained below the level implied by policy restrictions, suggesting delayed behavioural adjustment and limited initial responsiveness to policy signals. More broadly, the findings indicate that behavioural responses to pandemic policy are dynamic, shaped by uncertainty and behavioural inertia, and sensitive to the frequency and complexity of policy changes.

Keywords: policy effectiveness and uncertainty, COVID-19, individual risk behaviour

1 INTRODUCTION

According to Mathieu et al. (2020), as of 15 September 2024, North Macedonia, Bosnia and Herzegovina, Croatia, Slovenia and Montenegro were among the nine countries with the highest total confirmed COVID-19 deaths per 100,000 people globally. Among the former Yugoslavian countries, only Serbia and Kosovo were outliers, ranking 40th and 68th, respectively.

Given the high COVID-19 death toll in these countries, along with neighbouring Bulgaria, it is important to investigate the underlying health system weaknesses, social factors, and policy responses that may have contributed to these outcomes. Analysing these elements can also offer valuable insights for other policy areas – such as economic, social, and environmental policies – by highlighting systemic issues that affect broader aspects of governance and policy effectiveness.

Potter and Harries (2006) suggest that the "determinants of policy effectiveness" for health policy models require that public administration systems include diverse social, cultural, and economic motivators to encourage the desired behaviour – or "adherence" – to policy measures. The COVID-19 pandemic represents a rare opportunity to empirically analyse adherence behaviour during a rapidly changing crisis and corresponding responses to policies intended to minimise infection rates and protect public health. The literature on the COVID-19 pandemic spans three primary analytical domains: policy efficacy, policy determinants, and public adherence to policy measures. The first group of papers evaluates how non-pharmaceutical interventions (stringency measures), such as stay-at-home policies, lockdowns, and school closures, directly impacted infection rates and mortality (Hsiang et al., 2020; Walker et al., 2020; Fuller et al., 2021; Chen, Li and Wu, 2022).

In parallel, a second group of researchers investigates the political and economic motivations that led governments to adopt specific rules. These studies often find that a country's response was less a reflection of viral load and more a result of economic constraints, human capital, and/or political preferences (Capano et al., 2020; Desson et al., 2020; Chen et al., 2021; Jalloh et al., 2022; Toshkov, Carroll and Yesilkagit, 2022; Kuo and Wen, 2023; Sonora, 2023).

The third category focuses on public adherence and behavioural moderation, acknowledging that government mandates only work if people follow them. By analysing how factors like communication strategies, risk perception, institutional trust and socio-economic status influence compliance, this research explains how human agency serves as a critical moderator between a written policy and its actual success in the real world (Bargain and Aminjonov, 2020; Cairney and Wellstead, 2020; Berry et al., 2021; Pagliaro et al., 2021; Berg-Beckhoff et al., 2022; Tu, Chen and Mesler, 2023; Luciano, Crova and Canella, 2023 and Brunekreef et al., 2025).

A related strand of the literature integrates multiple dimensions of pandemic policy analysis within a single empirical framework, including the interaction between policy determinants and behavioural compliance (Waterschoot et al., 2023; Agyapon-Ntra and McSharry, 2023), policy formation and policy effectiveness (Cross, Ng and Scuffham, 2020; Herreros and Svendsen, 2024), and policy effectiveness and public adherence (Dave et al., 2021).

In the context of this study, adherence refers specifically to *mobility*-related behavioural responses captured through aggregate movement data, rather than other forms of pandemic compliance, such as masking or vaccinations. Because adherence to public health policy is dynamic and behaviourally mediated, static and panel-based empirical frameworks may obscure important temporal variation in residents' responses to government restrictions by averaging away behavioural adjustment within individual countries over time. We therefore employ a high-frequency time-series approach to capture the evolving relationship between policy stringency and mobility within individual nations. This framework enables us to analyse whether deviations from policy prescriptions represent gradual convergence towards policy objectives or short-run behavioural responses to uncertainty and perceived utility costs. Behavioural responses may also exhibit inertia, whereby individuals adjust only gradually to changing restrictions and information environments, generating persistent deviations between observed behaviour and policy objectives over time. More generally, these dynamics may be consistent with behavioural mechanisms, such as those associated with myopic loss aversion, policy fatigue, and rational inattention.

Following the work of Agyapon-Ntra and McSharry (2023), Luciano, Crova and Canella (2023) and Waterschoot et al. (2023), who utilised a wide range of control variables (such as GDP, literacy, and corruption), we employ structural break analysis to capture the mediating effects of diverse cultural and religious settings on adherence levels across the analysed countries.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

In addition to these behavioural considerations, the empirical application of our model focuses on a strategic subset of South-Eastern European countries (Croatia, Bosnia and Herzegovina, Slovenia and Serbia), characterized by weak yet heterogeneous cultural and institutional frameworks and significantly varied levels of public trust in government. This specific institutional context provides a useful setting for testing behavioural aspects of coherence, as it allows us to examine how varying cross-countries heterogeneities moderate the relationship between official mandates and actual compliance. Focusing on this regional context, we examine whether different environments help contextualise persistent deviations from policy adherence and whether behavioural responses differ systematically across institutional settings. More broadly, these dynamics may be consistent with a wide range of behavioural mechanisms identified in behavioural economics, such as attention and inertia.

At its core, the decision to remain at home during the pandemic reflects both behavioural responses to mobility restrictions intended to reduce viral transmission and unobserved idiosyncratic individual choice. During the pandemic, individual countries implemented mobility stringency policies to prevent the spread of COVID-19 among their population. The relationship between government stringency and public health outcomes is not direct, but instead depends critically on public adherence to policy measures. More precisely, it is moderated by "resident adherence" – the degree to which citizens follow those rules, which is a function of cultural norms.

Google's cell phone data allows us to track aggregate resident mobility and identify patterns of behavioural change: mobility should decrease as stringency increases (if policy is effective), and differences in preferences across the countries should lead to idiosyncratic responses to policy recommendations. Accordingly, our measure of behavioural adherence is limited to mobility-related responses to policy restrictions rather than broader dimensions of public health compliance. We study the properties of unobserved individual behaviour (adherence) using the estimated residuals from a model of individual mobility, which is conditional on policy restrictions and the level of adherence in each country in the sample. The estimated residual represents the unobserved mobility preferences of individuals and is an indicator of how closely individual and policymaker preferences align.

We are also interested in the time series properties of this estimated behavioural residual. Persistent deviations from official policy guidance therefore reflect divergence between government objectives and observed public behaviour. Residual deviations from expected mobility behaviour may reflect changing risk perceptions, behavioural inertia, and periods of under- or over-compliance relative to official policy guidance. We evaluate whether behavioural residuals gradually converge towards policy objectives and the time required for such adjustment. Mean reversion theory suggests that regret, fear, or risk will converge towards a normal behavioural state time over time. This raises the question of whether individual behaviour follows a stochastic process characterised by temporary deviations from, but eventual convergence towards, policy-consistent behaviour.

Using standard, structural break, and rolling unit root tests, we reject a unit root process if individual behaviour eventually converges as prescribed by the policy. Our estimates cannot determine whether individuals are risk-taking or risk-averse, as we only observe changes in risk attitudes relative to policy recommendations. Behavioural adjustment may be interpreted through the lens of a wide range of behavioural economics of attention and inertia theories (Thaler et al., 1997; Berg, Dickhaut and Rietz, 2010; Costa, 2018).

Overall, the analysis suggests that the policy is somewhat effective, as suggested by Hypothesis 1, but behavioural responses are dynamically inconsistent. Our estimates do demonstrate that residents in the short to medium run respond relatively "weakly" to the policy measures intended by health officials, but we find evidence of predicted compliance in the long run. We also find that in the medium run, residents adjust their behaviour further away from policy prescriptions – except for Serbia which exhibits a form of behavioural overshooting. Furthermore, our results suggest policy effectiveness has distinct regime shifts, as in Hypothesis 2, with structural breaks corresponding closely to divergence in religious and school calendars across the region.

Additionally, we find that behavioural inertia as well as shocks in individual preferences prevent both immediate and persistent full-scale adherence – as stated in Hypothesis 3. When individuals are faced with uncertainty regarding the intended or likely outcome of policy measures, observed responses remain anchored to established routines rather than adjusting fully to the new policy signal. The estimated idiosyncratic component may therefore partly reflect the degree to which uncertainty strengthened rather than weakened behavioural inertia, with the pull of prior routines intensifying precisely when the expected benefit of departing from them was unclear. The results nevertheless suggest gradual long-run convergence between resident behaviour and policy objectives.

2 TIMELINE OF THE COVID-19 PANDEMIC

Throughout the analysed period, the four countries in our sample experienced five to six major waves of COVID-19 cases (Mathieu et al., 2020). The first cases were reported in March 2020, and the countries in our sample promptly declared states of emergency or pandemic status, implementing restrictive measures. Slovenia and Croatia closed their borders with Italy and shut down all non-essential businesses. Additionally, Croatia imposed restrictions on inter-city travel (Hale, 2021; Lazić, Lazić and Kolarić, 2020). Moreover, all countries banned public gatherings, closed schools, and transitioned to online learning (Hale, 2021). Serbia took more extreme measures, introducing curfews and restricting the movement of adults 65 and over (Uzelac, 2023).

Shortly after the introduction of restrictions, the COVID-19 case curve flattened, leading to the gradual reopening of economies by late spring 2020 (Mathieu et al., 2020). For instance, Slovenia declared victory over the pandemic on 15 May 2020

(Vladisavljević, 2020). Croatia and Serbia soon followed, even holding elections in June and July 2020 (Karlović, 2021; Uzelac, 2023). As a tourism-dependent economy, Croatia relaxed restrictions to salvage the summer tourist season, reopening its borders to tourists, particularly from the EU (Vlada RH, 2020).

The timeline of the second wave of COVID-19 varied among countries, as did the responses of policymakers. Serbia and Bosnia and Herzegovina (hereafter B&H) were the first to experience the second wave, which began in July 2020 (Mathieu et al., 2020; Hale, 2021). In Serbia, this wave prompted the government to reintroduce restrictions, including curfews. These measures triggered large protests in cities such as Belgrade, where thousands demonstrated against the government's handling of the crisis, the reintroduction of curfews, and the perceived politicization of the pandemic. The protests turned violent, leading to clashes with the police (BBC, 2020).

Conversely, political fragmentation between the two federal units within B&H (the Federation (FBiH) and Republika Srpska (RS)) resulted in inconsistent measures across B&H – the RS imposed stricter curfews, while FBiH was slower to respond. The lack of a coordinated national strategy made it more difficult to control the spread of the virus (Milinković, 2021; Sajić, 2021).

In Slovenia and Croatia, the second resurgent wave occurred in the autumn of 2020, prompting the respective governments to reintroduce restrictions that remained elevated until the spring of 2021. A similar situation unfolded in B&H and Serbia in November and December 2021. Compared to the spring lockdown, the responses of all governments, except for Slovenia, were more moderate (Mathieu et al., 2020; Hale, 2021). In Slovenia, the government reinstated restrictions, including curfews, school closures, and limitations on social gatherings (Ružić Gorenjec et al., 2021).

The vaccine roll-out began in January 2021 in Croatia, Serbia, and Slovenia, while Bosnia and Herzegovina initially struggled to secure the vaccine (Mathieu et al., 2020). Unlike many EU countries, such as Croatia and Slovenia, B&H did not have direct access to the EU's vaccine procurement program and relied on the global COVAX initiative, which was slow in delivering doses (Fojnica et al., 2021). In contrast, Serbia's ability to secure a diverse range of vaccination supplies enabled it to inoculate its population rapidly and to offer vaccines to neighbouring countries (Dartford, 2021).

In the spring of 2021, all four countries experienced a surge in the number of cases attributed to the Alpha variant, and all but Bosnia and Herzegovina saw another surge driven by the Delta variant in the autumn of the same year. A significant surge occurred at the beginning of 2022 with the Omicron variant. The final surge took place in the summer of 2022, and was much milder compared to previous waves, except in Serbia, which recorded major increases in the number of cases (Mathieu et al., 2020).

Anti-COVID stringency policies declined throughout 2021 even as cases slowly increased during the spring and autumn waves. Slovenia maintained relatively high stringency levels throughout the year, while B&H showed the smallest variation in stringency. Croatia, due to its reliance on tourism, adopted a more laissez-faire approach starting in the summer months (Hale, 2021). In the summer of 2021, both Croatia and Slovenia introduced EU Digital Green Certificates, which were required for attending events, gatherings, and entering certain workplaces (European Commission, 2021; Healthcare IT News, 2021).

All countries experienced protests and significant political polarisation regarding policy responses. Serbia was the first to face protests in July 2020 due to the extreme level of stringency (BBC, 2020). In Slovenia, protests erupted in the autumn of 2020, with citizens expressing dissatisfaction with the government's handling of the crisis and restrictions on personal freedoms (Vladisavljević, 2021; Kriesi and Oana, 2023). In B&H, the slow vaccine roll-out and perceived mishandling of the pandemic led to public discontent in spring 2021, prompting protests in major cities where citizens demanded a more effective government response, including faster vaccination and more transparent decision-making (Reuters, 2021). In Croatia, large protests occurred in cities such as Zagreb and Split against the mandatory use of EU Digital Green Certificates for workplaces and events in autumn 2021, reflecting significant public opposition to vaccine mandates and restrictions (Grgurinović, 2021).

3 POLICY EFFECTIVENESS: THEORETICAL BACKGROUND

Given the previous overview of the pandemic policies and responses, we couch the discussion in terms of policy effectiveness, or adherence, and uncertainty. Whether a policy is effective is determined by how closely residents follow the prescribed restriction and requires the designer to understand the preferences of their constituents. Thus, as in Brainard (1967), we can consider policy effectiveness, for any time t, using the following relationship

$$B_t = \beta P_t + \eta_t, \ \beta \neq 0 \tag{1}$$

where individual target behaviour is represented by B, P is a vector of policies implemented, and the vector $\eta_t \sim iid\left(0, \sigma_\eta^2\right)$ represents other exogenous idiosyncratic factors influencing an individual's actions, "animal spirits".[1] Note that P and η are time-dependent, representing that new information may change both policy and individual policy responses. From the household's perspective $P_t \sim iid\left(0, \sigma_P^2\right)$. In this simple formulation, β is the residents' policy response parameter – representing resident preferences – and is known to individuals but unknown to the policymaker, and is the parameter of primary interest in this paper.

[1] For example, Hale (2021) identify four different types of COVID policies: economic support, containment and health, government stringency, and overall stringency.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

If any given policy is ignored $\beta = 0$. On the other hand, others may believe the policy is insufficient, in this case $|\beta| > 1$ which could be understood as residents' belief that the policy is inadequate to achieve its goals or that the policy does not fully account for risk. Aggregating over all individuals, it is reasonable to assume that the average policy response is given as $|\bar{\beta}| \leq 1$. In this context, deviations from the prescribed policy are given in η, which reflects individual's set of preferences, news and information shocks, and animal spirits. Moreover, η represents the individual's level of understanding of the policy at time t and reflects a lag between the policy announcement and the behavioural adjustment.

From the policymaker's perspective, policy is *effective*, regardless of whether it is the *correct* policy, if residents adopt behaviour required by the policy. The policymaker must design an optimal policy, P^*, based on any given policy response, say $\tilde{\beta}$, to achieve the policy goal, B^*,

$$B_t^* = \tilde{\beta} P_t^*$$

That is, the preferences of both the residents r and policymaker p are equal:

$$U_r(B^*) = U_p(P^* \mid \tilde{\beta}) \tag{2}$$

Policy "confusion", or uncertainty, is determined by the variability in either policy saliency, contradictory policies, and/or frequent policy revisions. Angelini et al. (2023) define their policy function evolving as:

$$P_t = \varphi P_{t-1} + (1-\varphi) P^*, |\varphi| \in (0,1) \tag{3}$$

where P^* is the policymaker's optimal response to minimising an economy-health loss function, as estimated in Sonora (2023). In this formulation, actual policy responses are staggered over time as the weighted average of previous policy choices, with policy inertia given by φ, and the current optimal policy. Note that as $\varphi \to 0$, policy converges to the optimal response, reducing uncertainty. Solving this recursively, we get

$$P_t = \varphi^t P_0 + \lambda \sum_{j=1}^{t} \varphi^{t-j} P_j^*$$

where $\lambda \equiv (1 - \varphi)$ is an adaptive expectations adjustment parameter. *If* we assume that policy is constant $P^* = P$ and substitute this into equation (1), after passing through the expectation operator, we have

$$B_t = \beta P_t \tag{4}$$

Now suppose that residents' adjusted behaviour is dictated by current policy and their own lagged behaviour, or

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

$$B_t = \beta P_t + \rho B_{t-1} + \eta_t, |\rho| \in (0,1) \tag{5}$$

The AR portion of this equation represents habit persistence and/or adjustment dynamics as residents update their responses to previous policy decisions. The behavioural adjustment in equation (4) becomes

$$B = \beta_{LR} P \tag{6}$$

where $\beta_{LR} = \frac{\beta}{1-\rho}$. Note that the relationship between β and β_{LR} is dependent on the sign of ρ; if $\rho \in (-1,0)$ then $\beta_{LR} < \beta$. In this case, residents reduce their policy compliance over time, in support of myopic loss aversion. If, however, $\rho \in (0,1)$ then residents' preferences converge to their short-term responses, which is the result of "present bias" (see O'Donoghue and Rabin, 1999), information constraints, and/or rational inattention as a form of "behavioural over-shooting", e.g. Maćkowiak and Wiederholt (2009).

Therefore, we expect that individual behavioural response parameter β will be dynamically inconsistent at different time lags as well as country-specific, as suggested by Hypothesis 1.

Hypothesis 1: Policy effectiveness in terms of alignment between mobility behaviour and policy recommendations is complex, dynamically inconsistent and incomplete in the short to medium run.

This analysis suggests that because $E(\eta)=0$ policy makers implicitly disregard idiosyncratic individual behaviour, irrespective of whether they know residents' true policy and behavioural response parameters. Moreover, these parameters are likely to create behaviour which deviates from policymakers' intent.

If we relax the assumption that policy is constant and substitute equation (3) into equation (1) and assume $E(P, \eta) \neq 0$ and P_t and P_t^* are time variants, and recognise that η is non-constant, the variance of equation (1) is given as, after rearranging to behavioural variance, or policy confusion, as

$$\sigma_B^2 = \beta^2 \frac{(1-\varphi)^2}{1-\varphi^2} \sigma_{P^*}^2 + \sigma_\eta^2 + 2\beta\sigma_{P\eta}$$

where σ_B^2 is the variance of variable i and σ_{ij} is the covariance of variables i and j. We assume $Cov(P_{t-1}, \eta_t) = 0$, which, given the "backward" nature of policy is reasonable, as previous policy is uncorrelated with unforeseen current behavioural shocks. In this formulation, as $\varphi \to 0$ there is a reduction in policy confusion. As a special case, suppose policy is always optimal, i.e. if $P_t^* = P_{t-1}$ then $Cov(P_t, P_t^*) = Cov(P_t, \eta_t) = 0$, and (7) collapses, after rearranging, to find the source of unobservable behavioural variance, to

$$\sigma_B^2 = \sigma_\eta^2$$

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

implying behavioural volatility is purely a function of animal spirits. Put another way, even without policy confusion, there will still be policy ineffectiveness due to idiosyncratic behaviour.

Following the same logic, if the behavioural function now includes the AR term, behavioural variance becomes

$$\sigma_B^2 = \frac{1}{\left(1-\rho^2\right)}\left[\beta^2\sigma_P^2\frac{1+\varphi\rho}{\left(1-\varphi\rho\right)}+\sigma_\eta^2+2\beta\sigma_{P\eta}\right]$$

As before, the policy uncertainty analogue, σ_B^2, to policy effectiveness, β, is dependent on the sign of $|\rho| \in (0,1)$. For $\rho<0$, behavioural uncertainty falls while the opposite is true if $\rho>0$.

Due to cross-national variation in religious and national settings, we expect that differences in national and religious calendars will have a moderating role in the mean-reverting properties of residuals η_t as suggested by Hypothesis 2.

Hypothesis 2: There is baseline cross-national variance in the rate and extent of compliance and convergence to policy, which is further compounded by localized structural breaks.

Furthermore, if agents' behaviour gradually adjusts to policy targets, the discrepancy between policy-prescribed outcomes and observed behavioural responses, η_t, should – at least in the long run – exhibit mean-reverting properties. That is, deviations from policymakers' intended path should be temporary and tend to converge back over time, as suggested by Hypothesis 3.

Hypothesis 3: Behavioural inertia and shocks in individual preferences prevent both immediate and persistent behaviour in an evolving crisis leading to uncertainty about policy effectiveness.

4 METHODOLOGY

To test three hypotheses within this theoretical framework, we employ a two-stage empirical strategy. First, we test Hypothesis 1 by noting that individual mobility can be decomposed into observed and unobserved components. Because some consumer behaviour is dictated by policies designed to restrict movements with the goal of minimising risk associated with COVID-19, this component can be estimated. Any behaviour that differs from that which is "dictated" by policy represents individual choice and can be estimated using the following ARDL specification:

$$MI_t = \alpha + \rho MI_{t-7} + \beta\left(L\right)OxSI_t + X_t\gamma + \eta_t \tag{7}$$

where *MI* is the mobility index, representing individual revealed preferences, and *OxSI*, for stringency index, is a country-specific policy variable. Lagged *MI* rep-

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

resents inertia in individual behaviour, from equation (5), and L is a lag operator. We use concurrent and 7- and 14-day lags of *OxSI* in our specification, with time between lags reducing collinearity in policy decisions. We include the one- and two-week lags because most COVID-19 data on cases and deaths were presented to the public as a rolling 7-day moving average and 14 days corresponds to the length of quarantine should a test be positive. The inclusion of lags also allows individuals to adjust to new policy interventions.

There might be concern that policy prescriptions and behaviour are endogenous. That is, policymakers respond to individual voter preferences when determining pandemic restrictions. However, this would require very frequent changes in policymaking behaviour. Moreover, Sonora (2023) demonstrates that the level of stringency is determined by minimising an "economy-health" loss function and can, therefore, be considered exogenous. X is a vector of other country-specific and national controls, discussed in the data section below. Similarly, Jalloh et al. (2022) showed that policy was driven by information on deaths, health expenditures as a percent of GDP while Sonora and Tica (2024) find no evidence for policy endogeneity in their study of four US states during the pandemic.

Of primary interest are the contemporaneous estimates behavioural elasticity with respect to policy, $\hat{\beta}_0$ and the longer-term adjustment elasticity estimates derived from $\hat{\rho}$ and $\hat{\beta}(L)$. These estimated coefficients represent the policy response parameters introduced in equations (1) and (5) and are an empirical test of Hypothesis 1 with the null given as no policy alignment. Given that people require some time to fully incorporate policy into their decisions, we anticipate the short-run estimate to be more inelastic than the longer-term elasticity. This reflects the benefits of using very high frequency data in empirical modelling; it allows us to better understand individual behaviour in the immediate short run and compare it to individual choices as the time horizon lengthens. In an urgent situation, such as a pandemic, the initial response to new restrictive policy is not likely to be as effective as policymakers may believe.

The second aspect of interest for understanding unobservable human behaviour with respect to policy intended to inhibit social interaction is the time-series properties of η which are the estimated residuals from equation (7). In this context, $\hat{\eta}$ represents country-level unobserved idiosyncratic behaviour and is the manifestation of individual preferences with respect to risk, politics, behavioural norms, information, etc. (see, for example, Allcott et al., 2020; Barrios and Hochberg, 2020; and Poletti, Ajelli and Merler, 2011).

To study the convergence of actual behaviour to that preferred by the policymaker, Hypotheses 2 and 3, we can apply unit root tests to the estimated residual, $\hat{\eta}$, which reflects unobservable individual choices. $\hat{\eta} \sim I(0)$ is interpreted as the revealed preferences of residents (r) converging over time to those of the policymakers (p) in country j

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

$$\lim_{t \to T} U_p^j \left(M^\tau \right) = U_r^j \left(M^\tau \right) \tag{8}$$

where M_τ is the target level of mobility, the behavioural empirical analogue to equation (2). Empirically, we can represent this as an *AR(1)* model

$$\hat{\eta}_t = \varphi \hat{\eta}_{t-1} + \nu_t \tag{9}$$

where, from equation (3), $\varphi = (1 - \lambda)$ is the policy "inertia" parameter.[2] If $\lambda \in (0,1)$, policy confusion is less prevalent and $\varphi \to 0$, which means that the deviation from the stated policy objective is short-lived. However, as $\lambda \to 0$ individuals are less policy compliant and $\hat{\eta}$ converges to a random walk, implying a greater degree of policy uncertainty or confusion.

There are two observations about the equation above. First, the preferences of both individuals and policymakers differ across countries. Second, the above equation does not imply that the policymaker is making optimal health policy to reduce the spread of the disease, but rather the decision is made based on their idiosyncratic set of preferences, as in Sonora (2023). We also interpret this preference matching within the context of the Tiebout (1956) hypothesis.

We employ three models that test for the presence of unit roots. The first is the standard augmented Dickey-Fuller test. Secondly, we employ the one- and two-break tests proposed by Clemente Lopez, Montañés and Reyes (1998) (CMR). Structural break unit root tests are useful when analysing time series that may be subject to deterministic breaks which could be misinterpreted as permanent stochastic processes that bias standard unit root tests towards nonstationarity, leading to size distortions. Structural breaks are employed to examine the role of cultural dimensions in explaining cross-national variance, as proposed in Hypothesis 2.

This model can handle both "long-run" innovative and "short-run" additive, or level, shocks. For our purposes, breaks in the data represent structural changes in individual behaviour as it relates to health policy and could therefore be either of these types of inflection points. Alternatively, innovations represent changes in behavioural trends. Note that both breaks could manifest in either less or more mobility as the social environment changes.

It is important to note that the unit root tests do not detect whether behaviour is overly cautious or heedless of policy, just whether or not $\hat{\eta} \sim I(0)$ and reverts to desired policy outcomes. For each test, we allowed for up to 14 days of lagged differenced terms to account for serial correlation. The suggested rule of thumb for choosing the maximum number of lagged data, suggested by Schwert (1989), is about 18 days, but we chose to use 14, as it corresponds to the number of days required for quarantine.

[2] Clearly this could be an *AR(k), k ≥ 1*, model, but we use an *AR(1)* representation to simplify the discussion.

PUBLIC SECTOR ECONOMICS 50 (3) 373-404 (2026)

The third unit root model employed is the rolling-window ADF test, which allows for the examination of time-varying stationarity properties in the data. By applying the test over successive sub-samples, we can detect periods when the series becomes non-stationary, thereby capturing the potential effects of policy uncertainty and other structural changes on the persistence of the variable, as suggested by Hypothesis 3.

5 DATA

Our data are daily and the sample is the two-year period from 17 February 2020 to 14 February 2022, which roughly coincides with the end of the Omicron variant spike in: B&H, Croatia, Serbia, and Slovenia. To proxy for individual mobility behaviour, we use available mobile phone data to observe movements. Google's Community Mobility Report (Google, 2021) indexes six different types of mobility gathered from mobile phone data, by country: grocery and pharmacy, retail and recreation, parks, residential, work, and transit. We use the daily mean of *five* of these indices to derive an overall index of mobility, the Google mobility index (*GMI*).[3] We exclude "park" mobility – defined as "...trends for places like national parks, public beaches, marinas, dog parks, plazas, and public gardens", see Google (2021) – for two reasons. Firstly, stringency measures are designed to *restrict* movement and with many areas being closed off, outdoor mobility will rise as there are few other locations to go outside the home. Secondly, from a data standpoint, the park index varies greatly across the countries. The index is defined as the percentage difference between the mobility on any given day based on pre-pandemic mobility, $GMI \in (-100\%,\ 100\%)$. This data has similarly been used in the context of the spread of COVID-19 by Sulyok and Walker (2020).

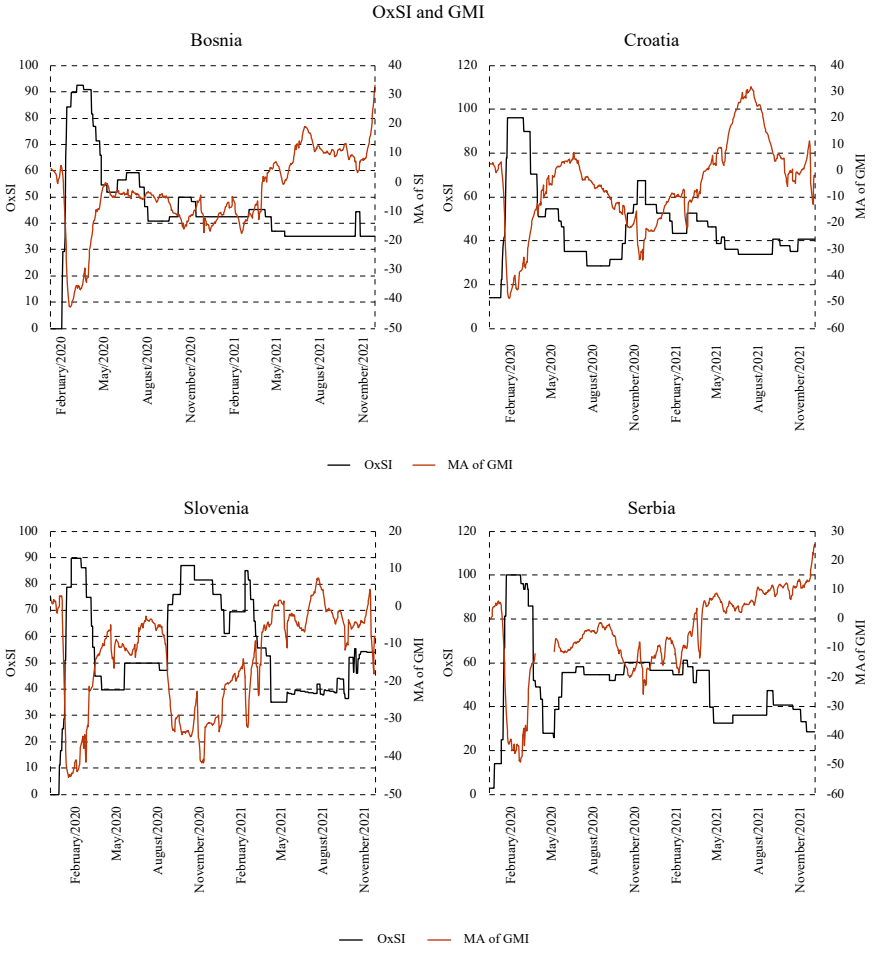
COVID-19 restriction policies are from the Oxford Coronavirus Government Response Tracker Stringency Index (denoted *OxSI*), see Hale (2021). This is one of several indices constructed by *OxSI* but was chosen as it quantifies the degree to which countries restrict individual movements, via lockdown policies, rather than economic support or health containment policies. The index is between 0 and 100 with a value of 100 being the most restrictive. The index represents all containment and closure policy indicators, such as school and workplace closing, public event cancellations, restrictions on gatherings, access to public transport, and travel, as well as stay-at-home recommendations. *OxSI* daily data captures the intensity and modality of policy interventions for COVID-19.

The control variables are national-level indicators. First, we include the number of new weekly cases or deaths per capita, available from Johns Hopkins (2021). This is used as an information variable; spikes in new recent cases indicate the spread of COVID-19 which leads to less mobility. Second, we include the current and one-week lagged percentage of eligible population that has been vaccinated, Vax rate, at least once. Finally, seasonal fixed effects are included for the summer period 1 June to 1 September for both 2020 and 2021.

[3] We also tried a principle components approach. Because the correlation between the mean and principle component index was over 0.94 for all countries we chose to use the mean as the calculation is more transparent.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

FIGURE 1
Mobility and stringency



Source: Authors.

Table 1 provides the mean, standard deviation (SD), the minimum and maximum of the *GMI* and *OxSI* for each country, in the top and bottom panels respectively. These data provide an overview of overall differences in policy and behavioural responses over the sample period. During the sample, Slovenia had the lowest average mobility, while the greatest degree of mobility volatility was in Croatia. This can be explained by the two extremes in Croatian mobility during this time, in particular the maximum amount of mobility which occurred during the summer tourist seasons in 2020 and 2021, as can be seen in figure 1, which shows both the *GMI* and *OxSI* for each of the countries. The increase was particularly acute in the summer of 2021. B&H and Serbia, and to a lesser extent Slovenia, follow similar patterns, exhibiting close to normal mobility until the first quarter of 2021 when mobility began to steadily increase.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

Table 1
Descriptive statistics

	Obs.	Mean	Std. dev.	Min	Max
Google Mobility Index (GMI)					
BIH	686	-3.7	13.8	-53.8	42.4
HRV	686	-5.1	18.0	-63.8	53.0
SLV	686	-14.0	16.1	-64.0	16.6
SRB	641	-4.8	15.5	-68.4	33.8
Oxford Stringency Index (OxSI)					
BIH	686	45.6	17.5	0.0	92.6
HRV	686	44.5	17.9	13.9	96.3
SLV	686	55.7	19.3	0.0	89.8
SRB	686	49.9	18.9	2.8	100.0

Note: GMI is in percent.
Source: Authors.

Part of the behaviour observed in figure 1 can be explained by considering the descriptive statistics for the *OxSI* in table 1. Overall, Croatia had the least restrictive policies while Slovenia had the most restrictive policy. However, Slovenia also had the largest variation in its policy, ranging from 0% to 89%. The only country with a 100% restrictive policy was Serbia, which occurred in the first quarter of 2020. We can see that both B&H and Serbia had similar policy patterns, while Croatia and Slovenia followed a different but similar pattern. Both of these countries undertook a considerable tightening of policy in late 2020 and early 2021 (see section 2 for more details on restrictions).

6 RESULTS AND DISCUSSION

Results for the model presented in equation (9) for each country can be found in table 2. Newey-West standard errors using 7 (one week) lags are in parenthesis, with *, ** and *** representing statistical significance at the 5%, 1%, and 0.1% levels, respectively. Estimates of most interest are at the top of the table 2. The first line, denoted $OxSI_t$, is the instantaneous, or day-to-day, response of behaviour to the current state of policy. The two-week response elasticity is calculated from the estimates of ρ and $\beta(L)$ and is denoted "*Response*". Mathematically, this is given as

$$Response = \frac{1+\hat{\beta}_7+\hat{\beta}_{14}}{1-\rho} \tag{10}$$

where the numeric subscripts refer to the number of lags. We also present the results for vaccination rate, weekly cases, and deaths. Finally, we include seasonal fixed effects to account for tourism impacts.

We begin our discussion by focusing on the estimated response coefficient rather than the immediate estimated elasticities. It is important to recall that the policy implemented by each country may not be optimal for reducing the spread of the virus; rather, it reflects the individual preferences of the policymaker's and trade-off between the economy and the health of the country's residents.

Table 2
Behaviour regression results – GMI dependent variable

	BIH	HRV	SLV	SRB	BIH	HRV	SLV	SRB
	GMI: using cases				GMI: using deaths			
OxSI	-0.437***	-0.510***	-0.568***	-0.482***	-0.430***	-0.504***	-0.543***	-0.484***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Response	-0.366***	-0.439***	-0.529***	-0.581***	-0.356***	-0.421***	-0.481***	-0.580***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GMI(-7)	0.596***	0.545***	0.537***	0.505***	0.596***	0.549***	0.545***	0.505***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
OxSI(-7)	0.051	0.179**	0.209***	0.080	0.048	0.179**	0.206***	0.081
	(0.415)	(0.011)	(0.007)	(0.138)	(0.452)	(0.010)	(0.008)	(0.131)
OxSI(-14)	0.238***	0.131***	0.114**	0.115***	0.238***	0.135***	0.118**	0.116***
	(0.000)	(0.000)	(0.017)	(0.001)	(0.000)	(0.000)	(0.016)	(0.001)
Vax rate	0.391***	0.189***	0.065***	0.121***	0.391***	0.172***	0.084***	0.125***
	(0.000)	(0.000)	(0.007)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Δ Cases	-0.000	-0.001***	0.001**	-0.000				
	(0.636)	(0.002)	(0.038)	(0.174)				
Δ Deaths					0.009	-0.049**	-0.017	-0.022
					(0.209)	(0.011)	(0.604)	(0.178)
Summer	3.388***	6.104***	1.839**	-2.217***	3.659***	6.109***	1.217	-2.334***
	(0.000)	(0.000)	(0.033)	(0.000)	(0.000)	(0.000)	(0.128)	(0.000)
R^2_a	0.869	0.826	0.712	0.825	0.869	0.825	0.711	0.825
F-stat	515.11	486.64	264.93	435.57	516.19	485.13	274.46	433.97

*Note: p-values in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01.*

Source: Authors.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

Interpreted through our behaviour-policy alignment premise (Hypothesis 1), the pattern that longer-run elasticities are typically smaller in absolute value than the instantaneous ones suggests that residents overreact on impact, but relax their response over the subsequent two-week period, which confirms Hypothesis 1 in terms of an incomplete and dynamically inconsistent response to policy – effectively mitigating the idiosyncratic component of the policy response, $\hat{\eta}$. Over time, that component regains influence, uncertainty rises, and the eventual response falls short of the initial adjustment. Both sets of estimates remain below the one-for-one benchmark, so the picture is one of incomplete compliance, with the key feature being decay from impact to the medium run rather than sustained alignment.

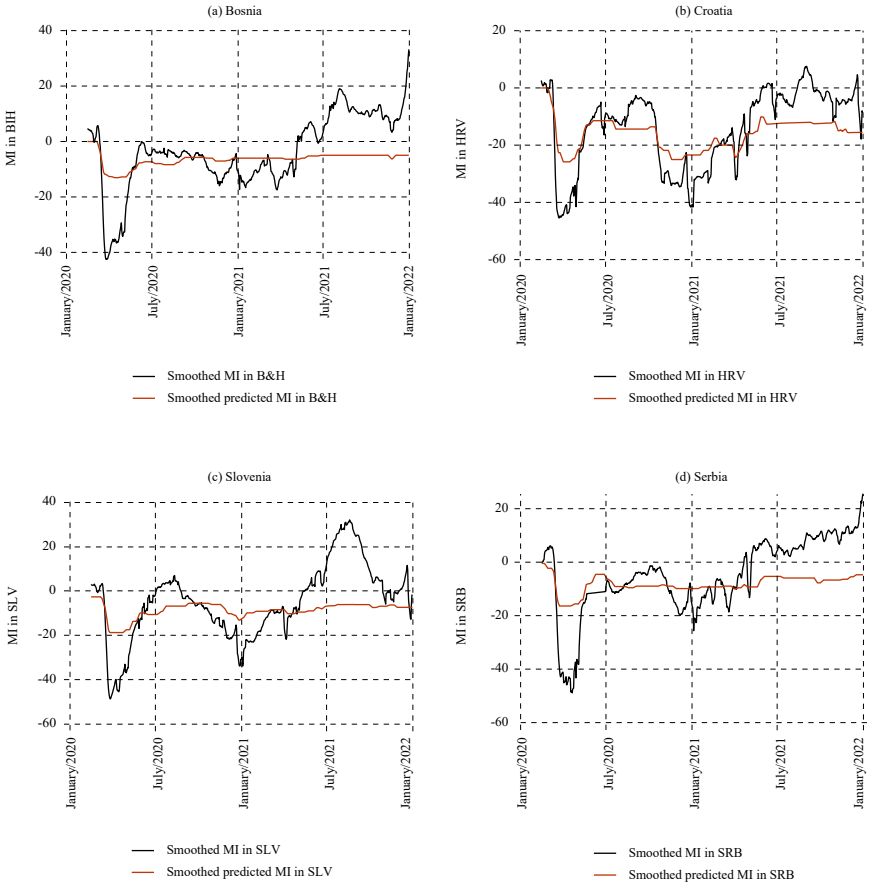
In Costa's (2018) study of the equity premium puzzle, he finds that higher inflation is associated with a lower equity premium, consistent with myopic loss aversion operating through short-horizon evaluation of outcomes. The results emphasise that when individuals focus on short-run fluctuations, loss-averse responses are amplified relative to a longer-run perspective. In the COVID context, this suggests that frequent, short-run updates on cases, deaths, or policy changes may have intensified behavioural responses compared to a setting where risks are evaluated over longer horizons.

The cross-country exceptions are informative, confirming Hypothesis 2. In Serbia, longer-run elasticities exceed the impact estimates (in absolute value), consistent with gradual adjustment, rational inattention, and/or present bias. Slovenia is the closest to the one-for-one benchmark on impact, yet the positive, significant coefficient on $OxSI(-14)$ points to a behavioural reversal over a two-week horizon – plausibly tied to quarantine dynamics or policy salience – again indicating that initial alignment is not fully maintained. Overall, these patterns reconcile negative, statistically significant elasticities with sub-unity magnitudes: compliance is immediate, incomplete, and not durably sustained.

A plausible behavioural mechanism for this pattern is myopic loss aversion (MLA). In this framework, after a policy is announced, individuals evaluate outcomes over short horizons and overweigh the perceived losses from falling short relative to the gains from exact compliance. This induces a precautionary "buffer"; behaviour adjusts sharply on impact but at a level that is difficult to sustain. As the period immediately after the announcement passes and expectations are better understood, the buffer attenuates, leaving a smaller longer-run elasticity. This result suggests that immediate responses are amplified, but over time, they lengthen the subsequent response time, potentially reducing policy effectiveness as the policy matures.

The positive $OxSI(-14)$ estimate for Slovenia is consistent with this overshoot-and-unwind dynamic, whereas Serbia's larger longer-run elasticities suggest gradual adjustment driven more by adaptation costs than by loss-averse buffering. In Serbia's case, MLA is "flipped". Initial behavioural responses are relatively small and then become more aligned to policy restrictions as more information is collected. Therefore, in Serbia's case, policy becomes more effective over longer time horizons.

FIGURE 2
Smoothed results

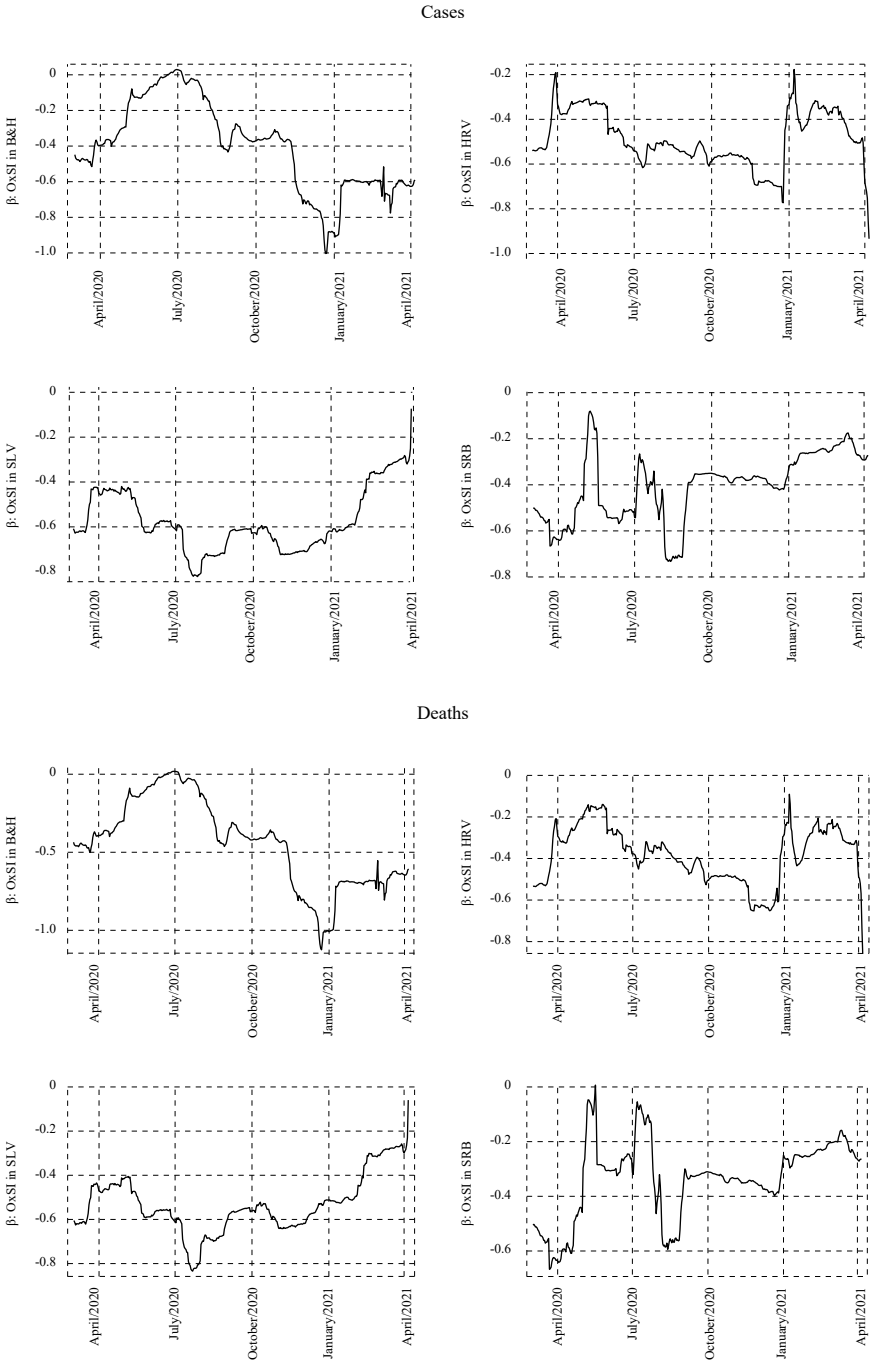


Source: Authors.

Figure 2 shows the smoothed relationship between estimated behaviour, that is behaviour explained by policy and actual behaviour in each of the countries – it is the difference between these two lines which represents the unobserved portion of behaviour. This illustrates policy response behaviour; the smaller the distance between these two lines, the greater its effectiveness. This figure illustrates that residents "overcompensated" their behaviour in the early stages of the pandemic, demonstrating a large degree of risk aversion as suggested by Hypothesis 1 and Hypothesis 3 in terms of uncertainty. During the middle stages, there appears to be a convergence of policy and behaviour. However, beginning in the first quarter of 2021, animal spirits took over and mobility expanded well beyond policymaker's prescriptions, except for Slovenian residents, whose behaviour remained *relatively* aligned with regulations – discussed below.

FIGURE 3

270 days rolling response elasticities using cases and deaths



Source: Authors.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

An alternative explanation could be that Serbians and Slovenians trust their government more than the other countries. According to the Wellcome Global Monitor (Welcome Trust, 2021), in 2020 57% of Serbians trusted their government while in B&H it was only 25%, for Croatia and Slovenia it was 42% and 48% respectively. These patterns may suggest that Serbian residents exhibited stronger medium-run behavioural alignment with policy measures after an initial adjustment period. Alternatively, in the aftermath of the demonstrations in Belgrade in July of 2020, Serbian policy makers reduced restrictions (see figure 1), and Serbs' behaviour loosened relative to the new more relaxed policies. Third, Serb and Slovenian policy measures may have been more closely aligned with prevailing resident preferences.

To answer the question about how individual preferences changed over the course of the pandemic, we conducted 270-day rolling regressions of equation (3) for each country, presented in figure 3 using both cases and deaths as the COVID-19 information proxy. Changes in policy elasticities reflect residents' changing attitudes as they become weary of restrictions and as they update their responses to the dynamics of the disease. For example, an overtightening of restrictions could lead to greater mobility as residents weigh the cost-benefit of the policy and decide the cost is too high vis-à-vis the benefit as suggested by Hypothesis 1 and 3.

We might expect the elasticities to be close to one during the initial stages of the pandemic. We do find that there is an immediate reaction to the policy, but, interestingly, within a matter of weeks residents increase their mobility beyond policymakers' preferences. The only exception is in Bosnia in late 2020 and early 2021, which roughly coincides with the Delta variant of the disease. It was not due to a policy adjustment, as shown in figure 1, but rather a change in Bosnian behavioural responses to the disease, both when identified using cases and deaths.

Next, we estimate whether individuals' preferences align with the desired behaviour dictated by policy stringency, that is, is $\hat{\eta} \sim I(0)$? Results of the unit root tests can be found in table 3. We present the Studentized t-statistics for each of the tests. For the break tests, we provide the estimated break dates. If no break dates are tabulated, no statistically significant break date was estimated. Beginning with ADF results, we see that, across the board, $\hat{\eta} \sim I(0)$ over the full sample suggesting that policy and individual behaviour *do* converge over the entire sample period, which is roughly 700 days.

TABLE 3

ADF and CMR one and two – break unit root tests

	BIH		HRV		SLV		SRB	
t_{ADF}	-4.175		-4.252		-6.875		-5.109	
t_{IO}	-4.799		-5.919		-7.327		-5.895	
Break date	NA	–	4-Apr-21	–	NA	–	19-Apr-20	–
t_{AO}	-4.71		-7.129		-6.676		-5.711	
Break date	6-May-21	–	3-Apr-21	–	NA	–	18-Apr-20	–
t_{IO}	-5.032		-9.444		-6.920		-6.311	
Break dates	6-May-21	11-May-21	4-Apr-21	11-Sep-21	NA	NA	19-Apr-20	NA
t_{AO}	-4.601		-8.523		-6.006		-5.854	
Break dates	5-May-21	10-May-21	3-Apr-21	10-Sep-21	NA	NA	18-Apr-20	NA

Note: NA represents structural break dates which are not statistically significant.

Critical values: (1%, 5%, 10%)
ADF: (-3.43, -2.86, -2.57)
IO 1 Break: (-5.49, -4.95, -4.68)
AO 1 Break: (-5.36, -4.80, -4.58)
IO 2 Break: (-6.41, -5.96, -5.70)
AO 2 Break: (-6.16, -5.59, -5.32)

Source: Authors.

PUBLIC SECTOR ECONOMICS
50 (3) 373-404 (2026)

Because we find evidence of convergence across the sample, the CMR tests are more helpful in identifying specific dates where individual behaviour changed. Interestingly, the single-break test identifies significant structural breaks only in Croatia, Bosnia, and Serbia, while the two-break test reveals that two significant breaks occur only in B&H and Croatia. Importantly, the estimated break dates correspond more closely to national and religious holidays than to major identifiable policy changes – as in Hypothesis 2. This suggests that culturally salient collective events may play an important role in moderating behavioural adherence dynamics. In Bosnia the breaks correspond to Labour Day (1 May), Julian calendar Easter (2 May 2021) and the end of Ramadan (12 May 2021), in Croatia the breaks fall on the Gregorian Easter Holiday (4 April 2021) and the first week of school (6-10 September 2021), while in Serbia, the break corresponds with Julian Easter on 19 April 2020.

This interpretation is broadly consistent with evidence from the COVID-19 literature suggesting that behavioural responses during the early stages of the pandemic were strongly shaped by uncertainty, fear, and voluntary behavioural adjustment prior to formal policy intervention (Alfaro et al., 2020; Czeisler et al., 2021; Goolsbee and Syverson, 2021). Such dynamics may also suggest that the gradual convergence towards policy prescriptions may reflect evolving risk perceptions and behavioural adaptation over time.

Alternatively, residents may *undershoot* their behavioural adjustment. This can come from either high short-term information costs or rational inattention (Sims, 2003; Maćkowiak and Wiederholt, 2009). In the rational inattention framework, too much information and/or its frequency leads individuals to selectively "ignore" new policies because they face an information processing constraint. Consequently, individuals allocate attention selectively towards the subset of policy information perceived as most relevant. In either case – over- or undershooting – residents are making "rational mistakes" in response to frequent policy updates.

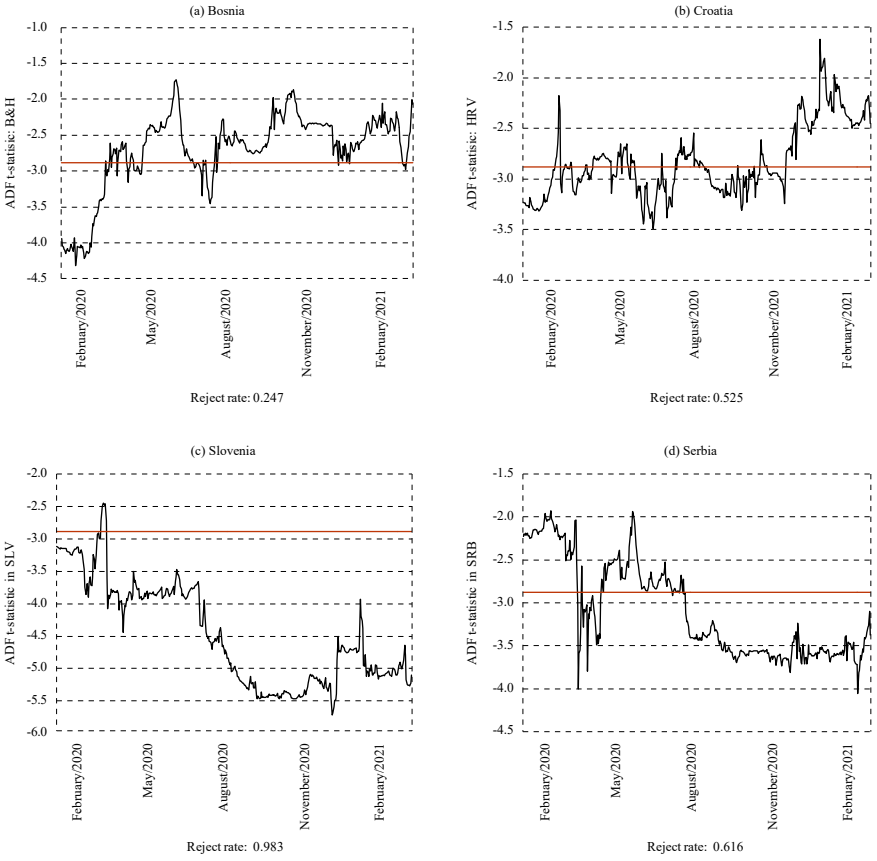
Differences in population density, infrastructure, and cultural "norms", as discussed in Best (2020), may contribute to cross-country differences in behavioural responses and policy outcomes in the region. Government responses also differed in both timing and levels of stringency, reflecting varying trade-offs between public health objectives and broader social, economic, and political considerations. Political incentives, voter preferences, and public opinion may also have influenced the formulation and implementation of pandemic policy across countries. Within our empirical framework, policy is treated as exogenously determined and reflects a trade-off between public health outcomes and broader economic costs.

6.1 DYNAMIC ALIGNMENT ANALYSIS

While these tests can identify specific dates of changes in behaviour, these tests do not necessarily let us observe the evolution of behaviour as new information or policy becomes available day-to-day. Put another way, does policy confusion persist over time and do frequent policy announcements exacerbate confusion, per Hypothesis 3? To investigate more changes in behaviour we conduct rolling Augmented Dickey-Fuller (ADF) unit root tests. We chose a rolling window of 270 days for each country and allow for a maximum of 14 days of lagged first differenced variables to control for serial correlation. The first window begins on 17 February 2020, and the last window begins on 21 May 2021 and should give us the necessary observations to have confidence in our estimates and to give sufficient time for individuals to adjust their behaviour either becoming more or less mobile than policy dictates. Also presented is the rejection rate of the null hypothesis of nonstationarity.

FIGURE 4

270 days rolling ADF tests



Note: Red line is 5% critical value.

Source: Authors.

Results of the rolling unit root tests are in figures 4(a) – 4(d). The figures show the $ADF - t$ statistic from each rolling unit root test and the horizontal red line is the 5% critical value, and again the first window is in February 2020. Periods when the $ADF - t$ falls below the critical value are interpreted as periods when individual behaviour conforms to each country's health policy – the shaded box is the percentage of the time in which individual mobility complies with policy recommendations relative to the entire sample, call this the "rate of compliance". We find that over a 270 days window, residents of Slovenia complied with state policymakers 76% of the sample period, particularly in the second half of 2020. On the other hand, Croatian residents only did so 13% of the time and only between late 2020 and early 2021. A possible explanation for the differences lies in the *Response* estimates presented in table 2. Based on the compliance rates, we might expect Croatian *Response* estimates to be closer to zero than in Slovenia, which is what we find when using both cases and deaths.

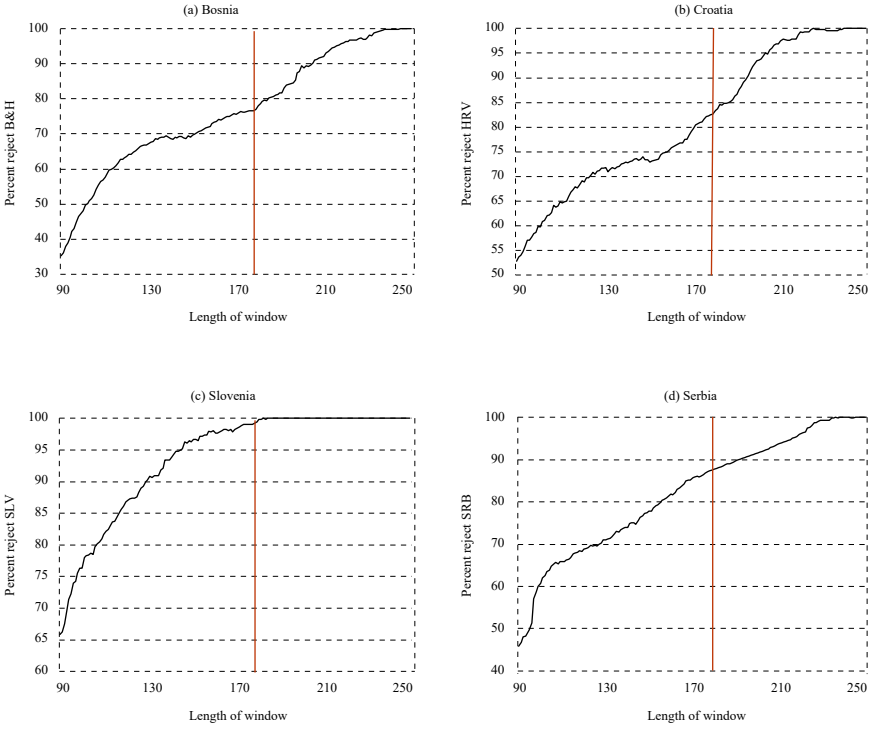
A second reason is that for the shorter time required to comply with policies is that policymakers may better understand their constituents. If voters believe that policy is either too lax or restrictive, their own behaviour will deviate from recommendations.

6.2 COMPLIANCE PERIOD

Given the relatively low percentage of behavioural convergence over the 270-day window, we next estimate the number of days required to achieve a near 100% compliance rate. As an additional test for Hypothesis 3, we conduct a range of rolling window unit root tests beginning with a 90-day window and ending with 250 days. The results can be found in figure 5 and demonstrate the percentage of each window length for which $\hat{\eta} \sim I(0)$. The vertical line indicates the 180-day rolling window. We find that residents of Slovenia meet a 100% compliance rate in about one-half of the year, followed by Croatia at 200 days. Both B&H and Serbia require an additional month or so to achieve full compliance. To put these results in perspective, using the same methodology Sonora and Gottwald-Belinić (2023) and Sonora (2024) find that US states and Western EU countries reach near 100% compliance in 365 and 225 days, respectively.

The observed slow convergence of behaviour towards benchmark responses suggests a persistent deviation from policy recommendations. While we have focused on an interpretation grounded in myopic loss aversion, whereby agents overweight recent policy signals and underweight accumulated information, alternative explanations may also account for relatively slow policy adherence dynamics. In particular, the patterns closely resemble those associated with pandemic fatigue and behavioural habituation. For example, using large-scale cross-country evidence, Petherick et al. (2021) document a gradual and widespread decline in adherence to high-cost protective behaviours during COVID-19, consistent with a form of fatigue in behavioural responses, alongside evidence of stabilisation and partial rebounds over time.

FIGURE 5
90 – 250 days rolling ADF rejection rate



Note: Red line at 180 days.
Source: Authors.

Complementary evidence using mobility data shows that the behavioural response to non-pharmaceutical interventions weakened over successive waves of the pandemic, suggesting a declining marginal responsiveness to repeated policy exposure (see, for example, Yilmazkuday, 2021). Taken together, this combination of attenuation and non-monotonic adjustment is consistent with a setting in which repeated exposure reduces behavioural responsiveness, even as new information continues to trigger short-run reactions. In this sense, the empirical patterns are consistent with both attention-based distortions, as emphasised by MLA, and a declining responsiveness arising from fatigue or habituation, implying that policy adherence may erode over time not only because agents misweigh information, but also because their responsiveness to that information diminishes with sustained exposure.

More broadly, the findings of this study align with several strands of the emerging literature on behavioural responses to pandemic policy. Consistent with Goolsbee and Syverson (2021) and Gupta et al. (2021), behavioural responses frequently diverged from the timing and intensity of formal policy interventions, suggesting that individual responses to uncertainty may evolve independently of official restrictions. Similarly, the observed attenuation in mobility responses over time

is broadly consistent with evidence on pandemic fatigue and a laxening of adherence to high-cost protective behaviours documented by Petherick et al. (2021) and Yilmazkuday (2021). At the same time, the presence of country-specific structural breaks associated with national and religious holidays suggests that socially and culturally salient events may play an important moderating role in shaping behavioural adherence dynamics across institutional settings.

These results have two implications. First, in a fast-moving crisis, such as a pandemic, policymakers may face delays in behavioural adjustment during rapidly evolving crises. Second, these findings suggest that simpler and less frequently adjusted policy measures may reduce behavioural confusion and improve longer-run adherence during periods of crisis.

7 SUMMARY

The COVID-19 pandemic has offered a unique lens through which to examine the intersection of public policy and individual behaviour in crisis settings. This study focused on mobility patterns and policy stringency in four former Yugoslavian countries – Bosnia and Herzegovina, Croatia, Serbia, and Slovenia – during the COVID-19 pandemic. By analysing mobility data and policy effectiveness, it became evident that the alignment of individual behaviour with policy recommendations is complex, occurring within differing social and institutional contexts.

The findings reveal that mobility behaviour does not immediately align with policy mandates. Initial responses showed significant over-compliance, driven by heightened anxiety and perceived risks. Over time, however, "animal spirits" emerged as individuals adapted, resulting in behaviours that diverged from prescribed measures. This divergence often reflects the evolving calculus of costs and benefits, with citizens weighing personal freedoms against public health concerns. Notably, this pattern was not uniform across countries, highlighting the critical role of localised social and political contexts.

Structural break tests and rolling regressions showed that behaviour eventually converges towards policy objectives over extended periods, though the rate and extent of compliance vary by country. Slovenia, for example, demonstrated relatively high adherence to policy, potentially consistent with higher institutional trust and differences in policy implementation. In contrast, Croatia exhibited lower compliance rates, which may be linked to its tourism-dependent economy and more lenient policy measures during key periods.

The analysis underscores the importance of designing policies that are not only clear and consistent but also aligned with public perceptions and societal preferences. Frequent policy shifts or overly stringent measures can erode public trust and lead to resistance, as observed in protests across the region. The findings also suggest that tensions between public health objectives and socio-economic considerations may influence long-run behavioural adherence.

Moreover, the study highlights the necessity of accounting for behavioural inertia and the time required for policy compliance to stabilise. In a rapidly evolving crisis, such as a pandemic, expectations of immediate compliance may be unrealistic. Instead, the results are broadly consistent with the view that simpler and clearer policy measures may reduce confusion and improve behavioural alignment over time.

The implications of this research extend beyond the pandemic context. Future crises, whether health-related or environmental, will similarly demand rapid policy responses and public cooperation. Lessons learned from the COVID-19 response in Southeast Europe emphasise the need for proactive communication strategies, trust-building efforts, and policies that reflect the lived realities of diverse populations.

Finally, this study contributes to the growing body of literature on policy effectiveness under uncertainty. By linking mobility behaviour to policy stringency, it provides a framework for assessing the dynamic interplay between public policy and individual decision-making. Future research could explore these dynamics in other contexts, examining how different governance structures, policy endogeneity, and cultural factors influence compliance and outcomes. Such insights will be vital as governments prepare for inevitable future crises.

While this paper focuses on a specific event, the COVID-19 pandemic, the results of this research can be easily applied to other similar negative shocks and may provide useful insights for future policy design during periods of crisis. Researchers are already looking for the next pandemic to hit, as seen in Yong (2022), and health institutes, such as the CDC, have begun reconsidering their approach to promulgating clearer, more concise, and more definitive policy to avoid miscommunication and improve the signal-to-noise ratio.

Disclosure statement

The authors have no conflicts of interest to declare.

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

During the preparation of this manuscript, Claude, Gemini and ChatGPT were used in the preparation of this document to format references, check the text for grammatical errors and typos, and improve the language clarity of key sections. The authors reviewed and edited any AI-generated content as needed and take full responsibility for the final manuscript.

REFERENCES

1. Agyapon-Ntra, K. and McSharry, P. E., 2023. A global analysis of the effectiveness of policy responses to COVID-19. *Scientific Reports*, 13(1), 5629. https://doi.org/10.1038/s41598-023-31709-2
2. Alfaro, L. [et al.], 2020. Social interactions in pandemics: Fear, altruism, and reciprocity. *NBER Working Paper*, No. 27134. https://doi.org/10.3386/w27134
3. Allcott, H. [et al.], 2020. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics,* 191, 104254. https://doi.org/10.1016/j.jpubeco.2020.104254
4. Angelini, E. [et al.], 2023. Modelling pandemic risks for policy analysis and forecasting. *Economic Modelling,* 120, 106162. https://doi.org/10.1016/j.econmod.2022.106162
5. Bargain, O. and Aminjonov, U., 2020. Trust and compliance to public health policies in times of COVID-19. *Journal of Public Economics*, 192, 104316. https://doi.org/10.1016/j.jpubeco.2020.104316
6. Barrios, J. M. and Hochberg, Y., 2020. Risk perception through the lens of politics in the time of the COVID-19 pandemic. *NBER Working Paper*, No. 27008. https://doi.org/10.3386/w27008
7. BBC News, 2020. Coronavirus: Belgrade protesters storm Serb parliament over curfew. *BBC News*, 8 July.
8. Berg, J. E., Dickhaut, J. W. and Rietz, T. A., 2010. Preference reversals: The impact of truth-revealing monetary incentives. *Games and Economic Behavior,* 68(2), pp. 443-468. https://doi.org/10.1016/j.geb.2009.07.008
9. Berg-Beckhoff, G. [et al.], 2022. Political stringency, infection rates, and higher education students' adherence to government measures in the Nordic countries and the UK during the first wave of the COVID-19 outbreak. *Preventive Medicine*, 164, 107245. https://doi.org/10.1016/j.ypmed.2022.107245
10. Berry, C. R. [et al.], 2021. Evaluating the effects of shelter-in-place policies during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(15), e2019706118. https://doi.org/10.1073/pnas.2019706118
11. Best, J., 2020. The social problems of COVID-19. In: G. W. Muschert [et al.], eds. *Social problems in the age of COVID-19: US perspectives*. Bristol: Bristol University Press, pp. 134-142. https://doi.org/10.56687/9781447359869-016
12. Brainard, W. C., 1967. Uncertainty and the effectiveness of policy. *The American Economic Review,* 57(2), pp. 411-425.
13. Brunekreef, J. [et al.], 2025. Impact of adherence and stringency on the effectiveness of lockdown measures: A modelling study. *PLOS One*, 20(12), e0338818. https://doi.org/10.1371/journal.pone.0338818
14. Cairney, P. and Wellstead, A., 2020. COVID-19: Effective policymaking depends on trust in experts, politicians, and the public. *Policy Design and Practice*, 4(1), pp. 1-14. https://doi.org/10.1080/25741292.2020.1837466
15. Capano, G. [et al.], 2020. Mobilizing Policy (In)Capacity to Fight COVID-19: Understanding Variations in State Responses. *Policy and Society*, 39(3), pp. 285-308. https://doi.org/10.1080/14494035.2020.1787628

16. Chen, D., Li, Y. and Wu, J., 2022. Policy Stringency, Political Conditions, and Public Performances of Pandemic Control: An International Comparison. *Public Performance & Management Review*, 45(4), pp. 916-939. https://doi.org/10.1080/15309576.2022.2040548
17. Chen, S. [et al.], 2021. Effective COVID-19 Control: A Comparative Analysis of the Stringency and Timeliness of Government Responses in Asia. *International Journal of Environmental Research and Public Health*, 18(16), 8686. https://doi.org/10.3390/ijerph18168686
18. Clemente Lopez, J., Montañés, A. and Reyes, M., 1998. Testing for a unit root in variables with a double change in the mean. *Economics Letters,* 59(2), pp. 175-182. https://doi.org/10.1016/S0165-1765(98)00052-4
19. Costa, R. B., 2018. A Systematic Test for Myopic Loss Aversion Theory. *Review of Behavioral Finance,* 10(4), pp. 320-335. https://doi.org/10.1108/RBF-06-2017-0059
20. Cross, M., Ng, S.-K. and Scuffham, P., 2020. Trading Health for Wealth: The Effect of COVID-19 Response Stringency. *International Journal of Environmental Research and Public Health*, 17(23), 8725. https://doi.org/10.3390/ijerph17238725
21. Czeisler, M. E. [et al.], 2021. Early public adherence with and support for stay-at-home COVID-19 mitigation strategies despite adverse life impact: A transnational cross-sectional survey study in the United States and Australia. *BMC Public Health,* 21, 503. https://doi.org/10.1186/s12889-021-10410-x
22. Dartford, K., 2021. Serbia inoculates neighbours as other Balkan countries receive doses of China's vaccine. *Euronews*, 28 March.
23. Dave, D. [et al.], 2021. When Do Shelter-in-Place Orders Fight Covid-19 Best? Policy Heterogeneity Across States and Adoption Time. *Economic Inquiry*, 59(1), pp. 29-52. https://doi.org/10.1111/ecin.12944
24. Desson, Z. [et al.], 2020. An analysis of the policy responses to the COVID-19 pandemic in France, Belgium, and Canada. *Health Policy and Technology*, 9(4), pp. 430-446. https://doi.org/10.1016/j.hlpt.2020.09.002
25. European Commission, 2021. EU Digital COVID Certificate: EU Gateway goes live with seven countries one month ahead of deadline. *European Commission Press Release*.
26. Fojnica, A. [et al.], 2021. Lack of Lockdown, Open Borders, and No Vaccination in Sight: Is Bosnia and Herzegovina a Control Group? *Preprint, Public and Global Health*. https://doi.org/10.1101/2021.03.01.21252700
27. Fuller, J. A. [et al.], 2021. Mitigation Policies and COVID-19 – Associated Mortality – 37 European Countries, January 23 – June 30, 2020. *MMWR. Morbidity and Mortality Weekly Report*, 70(2), pp. 58-62. https://doi.org/10.15585/mmwr.mm7002e4
28. Google, 2021. *Community mobility reports*.
29. Goolsbee, A. and Syverson, C., 2021. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics,* 193, 104311. https://doi.org/10.1016/j.jpubeco.2020.104311

30. Grgurinović, M., 2021. Mandatory COVID-19 passes cause discontent in Croatia. *Balkan Insight*, 23 November.
31. Gupta, S. [et al.], 2021. Tracking public and private responses to the COVID-19 epidemic: Evidence from state and local government actions. *American Journal of Health Economics,* 7(4). https://doi.org/10.1086/716197
32. Hale, T., 2021. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour,* 5(4), pp. 529-538. https://doi.org/10.1038/s41562-021-01079-8
33. Healthcare IT News, 2021. *Slovenia launches EU Digital COVID Certificate for its two million citizens*.
34. Herreros, F. and Svendsen, G. T., 2024. Social trust and the stringency of public policies against the COVID-19 pandemic. *European Politics and Society*, 25(5), pp. 866-885. https://doi.org/10.1080/23745118.2024.2331156
35. Hsiang, S. [et al.], 2020. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 584(7820), pp. 262-267. https://doi.org/10.1038/s41586-020-2404-8
36. Jalloh, M. F. [et al.], 2022. Drivers of COVID-19 policy stringency in 175 countries and territories: COVID-19 cases and deaths, gross domestic products per capita, and health expenditures. *Journal of Global Health*, 12, 05049. https://doi.org/10.7189/jogh.12.05049
37. Johns Hopkins University, 2021. *Coronavirus Resource Center*.
38. Karlović, A., 2020. Croatia 2020: An unexpectedly smooth victory for the incumbent. *Who Governs Europe?* 6 July.
39. Kriesi, H. and Oana, I.-E., 2023. Protest in unlikely times: Dynamics of collective mobilization in Europe during the COVID-19 crisis. *Journal of European Public Policy,* 30(4), pp. 740-765. https://doi.org/10.1080/13501763.2022.2140819
40. Kuo, F.-Y. and Wen, T.-H., 2023. Assessing the impact of local context and priorities regarding domestic disease outbreaks and imported risk on early pandemic response: Cross-continental comparisons. *Frontiers in Public Health*, 11, 1147768. https://doi.org/10.3389/fpubh.2023.1147768
41. Lazić, N., Lazić, V. and Kolarić, B., 2020. Prva tri mjeseca COVID-19 u Hrvatskoj, Sloveniji, Srbiji i Federaciji Bosne i Hercegovine – komparativna analiza mjera za suzbijanje bolesti. *Infektološki glasnik,* 40(2), pp. 43-49. https://doi.org/10.37797/ig.40.2.1
42. Luciano, F., Crova, F. and Canella, F., 2023. COVID-19 Containment Measures – a Step Back for Walking Mobility? A 2-Year, 60-Country Analysis of the Apple Mobility Data. *Journal of Physical Activity and Health*, 20(5), pp. 394-401. https://doi.org/10.1123/jpah.2022-0189
43. Maćkowiak, B. and Wiederholt, M., 2009. Optimal Sticky Prices under Rational Inattention. *American Economic Review,* 99(3), pp. 769-803. https://doi.org/10.1257/aer.99.3.769
44. Mathieu, E. [et al.], 2020. COVID-19 Pandemic. *Our World in Data*.
45. Milinković, I., 2021. Extraordinary measures in extraordinary times: Legal response to the COVID-19 crisis in Bosnia and Herzegovina. *Medicine, Law & Society,* 14(2), pp. 439-456. https://doi.org/10.18690/mls.14.2.443-460.2021

46. O'Donoghue, T. and Rabin, M., 1999. Doing it now or later. *The American Economic Review,* 89(1), pp. 103-124. https://doi.org/10.1257/aer.89.1.103
47. Pagliaro, S. [et al.], 2021. Trust predicts COVID-19 prescribed and discretionary behavioral intentions in 23 countries. *PLOS ONE*, 16(3), e0248334. https://doi.org/10.1371/journal.pone.0248334
48. Petherick, A. [et al.], 2021. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. *Nature Human Behaviour,* 5(9), pp. 1145-1160. https://doi.org/10.1038/s41562-021-01181-x
49. Poletti, P., Ajelli, M. and Merler, S., 2011. The effect of risk perception on the 2009 H1N1 pandemic influenza dynamics. *PLOS ONE,* 6(2). https://doi.org/10.1371/journal.pone.0016460
50. Potter, C. C. and Harries, J., 2006. The determinants of policy effectiveness. *Bulletin of the World Health Organization,* 84(11), p. 843.
51. Reuters, 2021. Bosnians protest, telling government to resign over lack of COVID vaccines. *Reuters*, 6 April.
52. Ružić Gorenjec, N. [et al.], 2021. COVID-19 in Slovenia, from a success story to disaster: What lessons can be learned? *Life,* 11(10), 1045. https://doi.org/10.3390/life11101045
53. Sajić, N., 2021. Federal institutional design and the COVID-19 crisis management in Bosnia and Herzegovina. In: I. Miljenović and E. Hrelja, eds. *Federalism and the response to COVID-19: A comparative analysis*. Routledge, pp. 49-58.
54. Schwert, G. W., 1989. Tests for unit roots: A Monte Carlo investigation. *Journal of Business & Economic Statistics,* 7(2), pp. 147-160. https://doi.org/10.1080/07350015.1989.10509723
55. Sims, C. A., 2003. Implications of rational inattention. *Journal of Monetary Economics,* 50(3), pp. 665-690. https://doi.org/10.1016/S0304-3932(03)00029-1
56. Sonora, R. J. and Gottwald-Belinić, M., 2023. Does individual behavior converge to policy recommendations in times of pandemic? Evidence from COVID-19 in US states. *EFZG Working Paper Series*, No. 23-04.
57. Sonora, R. J. and Tica, J., 2024. Individual behavior and policy response in times of a pandemic. *EFZG Working Paper Series*, 24-02.
58. Sonora, R. J., 2023. Estimates of a 'Pandemic Taylor rule' loss function. *Applied Economics Letters*, 30(15), pp. 2040-2045. https://doi.org/10.1080/13504851.2022.2089619
59. Sonora, R. J., 2024. *Household behavioral response and clubs to lockdown policy in Europe: Evidence from COVID*. Seminar presentation. Zagreb: Institute of Public Finance.
60. Sulyok, M. and Walker, M., 2020. Community movement and COVID-19: A global study using Google's Community Mobility Reports. *Epidemiology and Infection,* 148, e284. https://doi.org/10.1017/S0950268820002757
61. Thaler, R. H. [et al.], 1997. The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics,* 112(2), pp. 647-661. https://doi.org/10.1162/003355397555226

62. Tiebout, C., 1956. A pure theory of local expenditures. *Journal of Political Economy*, 64(5), pp. 416-424. https://doi.org/10.1086/257839
63. Toshkov, D., Carroll, B. and Yesilkagit, K., 2022. Government capacity, societal trust or party preferences: What accounts for the variety of national policy responses to the COVID-19 pandemic in Europe? *Journal of European Public Policy*, 29(7), pp. 1009-1028. https://doi.org/10.1080/13501763.2021.1928270
64. Tu, K., Chen, S. and Mesler, R. M., 2023. Policy stringency and the spread of COVID-19: The moderating role of culture and its implications on first responses. *Health Policy*, 137, 104896104896. https://doi.org/10.1016/j.healthpol.2023.104896
65. Uzelac, M., 2023. Using the pandemic for their own gain: The experiences of COVID in Serbia. *Anthropologica,* 65(1), pp. 1-23. https://doi.org/10.18357/anthropologica65120231066
66. Vlada RH, 2020. Božinović: Croatia is safe and one of few countries that will have tourism season. *Vlada RH*, 17 June.
67. Vladisavljević, A., 2020. Slovenia declares victory over coronavirus pandemic. *Balkan Insight*, 15 May.
68. Vladisavljević, A., 2021. Slovenians mobilise for another anti-Jansa push, demanding elections. *Balkan Insight*, 28 May.
69. Walker, P. G. T. [et al.], 2020. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science*, 369(6502), pp. 413-422. https://doi.org/10.1126/science.abc0035
70. Waterschoot, J. [et al.], 2023. How the Stringency of the COVID-19 Restrictions Influences Motivation for Adherence and Well-Being: The Critical Role of Proportionality. *International Journal of Health Policy and Management*, 12(1), pp. 1-13. https://doi.org/10.34172/ijhpm.2023.8021
71. Wellcome Trust, 2021. *Wellcome Global Monitor 2020: COVID-19*.
72. Yilmazkuday, H., 2021. Stay-at-home works to fight against COVID-19: International evidence from Google mobility data. *Journal of Human Behavior in the Social Environment*, 31(1-4), pp. 210-220. https://doi.org/10.1080/10911359.2020.1845903
73. Yong, E., 2022. The pandemic's legacy is already clear: All of this will happen again. *The Atlantic*, 30 September.

Transparent budget: the idea of deep transparency

DUŠAN PAVLOVIĆ, Ph.D.*
ŽIVKO BOJOVIĆ, Ph.D.*
VOJIN ŠENK, Ph.D.*

Article**
JEL: H61, H83, H57, D73, D82
https://doi.org/10.3326/pse.50.3.4

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

** Received: February 16, 2026
Accepted: May 13, 2026

Dušan PAVLOVIĆ
University of Belgrade, Faculty of Political Science, Jove Ilića 165, 11040 Beograd, Serbia
e-mail: dusan.pavlovic@fpn.bg.ac.rs
ORCiD: 0000-0003-3941-1539

Živko BOJOVIĆ
Singidunum University, Danijelova 32, 11010 Beograd, Serbia
e-mail: zivkocbojovic67@gmail.com
ORCiD: 0000-0001-6257-6417

Vojin ŠENK
University of Novi Sad, Faculty of Engineering Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
e-mail: vojin_senk@uns.ac.rs
ORCiD: 0000-0003-3029-0486

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

Abstract

Governments know exactly where public money goes. Citizens do not. This asymmetry is not a technical accident – it is a structural feature of how public financial management systems are designed. Existing transparency frameworks – the IMF Fiscal Transparency Code, the Open Budget Index, and the OECD Budget Transparency Toolkit – measure what governments choose to disclose. They do not make spending verifiable. This paper develops a model of deep transparency that closes this gap. By connecting budget allocations to real-time execution data through tamper-resistant technologies – write-once, read-many (WORM) storage and blockchain – the model enables citizens to trace every public outlay to its final private recipient. Deep transparency shifts fiscal accountability from publication to verification. It reduces the discretionary space through which patronage, corruption, and state capture operate. The model is normative and conceptual, but the problem it addresses is real and the tools required to solve it already exist.

Keywords: budget transparency, democratic accountability, public policy, digital governance, discretionary spending, WORM technology

1 INTRODUCTION

The idea that public budgets are inseparable from democratic accountability is not new. Aaron Wildavsky (1964: 1) famously described the budget as "the lifeblood of the government, the financial reflection of what the government intends to do" – the central arena where the political question of "who gets what and how" is answered. For Wildavsky, the budget was not a technical document but a political one, and the ability of citizens to scrutinize it was fundamental to holding governments accountable. That insight is now over six decades old. And yet, despite the internet, open data portals, and sophisticated public financial management systems, the promise it contained remains unfulfilled. Since the IMF introduced its Fiscal Transparency Code in 1998, governments have promised to disclose budget information. The promise has largely been kept – on paper. Budgets are published, procurement portals exist, and international organizations produce annual rankings of fiscal openness. Yet a critical gap persists: while governments maintain detailed records of every transaction internally, citizens lack the means to verify how public funds are actually spent. This asymmetry is not incidental. It is structural. And it undermines democratic accountability at its core.

The problem is most acute where it matters most. Research consistently finds that discretionary spending – including off-budget items, subsidies, and grants – generates systematic opportunities for abuse. Public procurement and public-private partnerships, now accounting for up to 33 percent of global GDP, are increasingly pre-rigged to favour politically connected firms (Fazekas et al., 2024), including in developed economies (Fazekas and Tóth, 2016; Dávid-Barrett and Fazekas, 2020; Curry, 2022). The result is corruption, state capture, party patronage, and rent-seeking that existing frameworks have failed to constrain (Rubin, 2020). They have failed not because the standards are wrong, but because they depend on the government's willingness to disclose.

Deep transparency addresses this democratic deficit. It is not primarily a technical upgrade to existing public financial management systems – those systems already exist, and governments already possess complete transaction-level data. Rather, deep transparency is a policy innovation that makes this data continuously verifiable by the public through tamper-resistant technologies such as write-once, read-many storage (WORM). By converting the insider advantage into public visibility, deep transparency restores the democratic principle that citizens have a right to scrutinize how their tax money is spent.

Existing public financial management systems – Integrated Financial Management Information Systems (IFMIS), International Public Sector Accounting Standards (IPSAS), the IMF Fiscal Transparency Code (IMF, 2019), and the IMF Digital Solutions Guidelines (IMF, 2023) – are designed for insiders. Government officials have full visibility over budget execution; citizens do not. Data can be modified locally, reports are generated episodically, and disclosure depends entirely on the government's choice. Why these systems fall short of democratic accountability and how deep transparency differs from them are developed in section 3.

This article is conceived as a normative, design-oriented framework rather than as an empirical case study or pilot evaluation. Its contribution lies in specifying the institutional logic, technical architecture, and implementation pathway through which deep transparency could be integrated into existing treasury and IFMIS-based public financial management systems.

The next sections proceed as follows. Section 2 develops the conceptual foundations of deep transparency. Section 3 situates the model within existing transparency practices and explains why current approaches remain insufficient. Section 4 presents the technical architecture that operationalizes the model in practice. Section 5 discusses the broader political economy implications of adopting deep transparency, and section 6 concludes.

2 DEEP TRANSPARENCY: THE MODEL

Our model presents a natural progression from previous models of transparency. We build on the arguments for openness, clarity in rules and decision-making, public scrutiny of government and business operations, and the application of transparency in budgetary and fiscal domains (Hood, 2006). According to Heald (2003), transparency entails that public expenditures are visible and understandable to designated user communities. This perspective underscores the inadequacy of simply generating and disseminating information; instead, transparency demands an audience capable of comprehending and effectively utilizing this information. It underlines the necessity of communicating public expenditure in a format intelligible to external stakeholders (referred to as "inward transparency"). This approach mirrors the accountability mechanisms expected of managers within publicly listed companies, where substantial resources are allocated to establishing accounting standards, financial reporting, and auditing conducted by accredited auditors (Heald, 2012).

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

It is important to note that deep transparency, as developed in this paper, is a normative and conceptual model. No government currently applies it in the form we describe, which is precisely why we wrote this paper.

The historical development of budget transparency can be divided into four stages, following the framework found in Khagram, Fung and de Renzio (2013):

- TS1 Absence of transparency (Weberian Model) – budgets not revealed to citizens.
- TS2 Paper transparency (new public management, NPM) – budgets revealed on paper, typically with lump sums (pre-digital era).
- TS3 Online transparency (new public governance, NPG) – data on public spending can easily be accessed online and are more detailed, categorized by economic classifications.
- TS4 Deep transparency – citizens can access the entire budget proposal process, adoption, execution, and associated discretionary spending, such as grants, subsidies, loans, and public contracts.

These four stages represent our analytical reading of how budget transparency has evolved historically, broadly corresponding to the trajectory documented in Khagram, Fung and de Renzio (2013). Deep transparency constitutes the fourth and most recent stage in this progression.

We define deep transparency as follows. It is a comprehensive model of budget oversight that enables real-time, detailed institutional tracking of public expenditures, linking each budgetary outlay to its final beneficiary, the recipient of the disbursed funds, including private agents in public-private partnerships. It employs write-once, read-many (WORM) storage and/or blockchain technology to ensure data integrity and prevent tampering, creating an immutable, publicly verifiable record of budget execution. It is essentially a part of financial technology, applying innovative technology-driven solutions to improve and automate the delivery of financial services. By making each stage of budget execution traceable and verifiable, the model significantly expands the scope and precision of fiscal disclosure.

To understand what deep transparency adds to the existing infrastructure, it is necessary to be precise about what existing systems already do – and where they stop. Governments already record detailed transaction-level data internally through treasury and IFMIS systems. This information is, in principle, comprehensive. In practice, however, it is not made publicly accessible. The information released externally is aggregated – typically at the level of programs, sectors, or institutions – rather than individual transactions. Reporting is also periodic: monthly, quarterly, or annually, rather than continuous or in real time. As a result, traceability is incomplete. It is difficult to follow specific budget outlays to their final beneficiaries, especially in areas such as public procurement or public-private partnerships. Deep transparency builds directly on these existing systems. It does not replace them. It extends them by making their underlying transaction data publicly accessible, traceable, and verifiable in real time.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

Our model develops the fourth stage of transparency by bringing the entire budget cycle into view and illuminating what we call the shadow budget. We define it as the broad domain of public spending that falls outside or at the margins of core budget documents and therefore escapes routine public scrutiny. It includes off-budget items such as contingency funds, special accounts, and quasi-fiscal activities of state-owned enterprises; tax expenditures such as exemptions and tax breaks that reduce revenues without appearing as direct outlays; debt obligations and guarantees; and discretionary transfers such as subsidies, grants, and loans to private and public entities (IBP, 2011). Public-private partnerships (PPPs) occupy a particularly opaque corner of the shadow budget: their costs are often spread over time or kept off the official balance sheet, concealing fiscal commitments that can rival conventional public debt in scale (IMF, 2019). Unlike the formal budget, where allocations are at least nominally visible, the shadow budget is fragmented – its components are scattered across separate platforms, agencies, and reporting frameworks, making it structurally difficult for non-insiders to follow. The shadow budget plays a significant role in generating corruption, state capture, rent-seeking, and party patronage, and in eroding public trust in political institutions (Bauhr et al., 2020).

Our model integrates subsidies, grants, loans, public procurement, PPPs, and other discretionary flows with the formal budget, exposing the areas most vulnerable to opacity and abuse. Every budget expenditure ultimately reaches a final recipient in the private sector, whether through a PPP contract, a subsidy, or any other budgetary channel. Our solution provides a straightforward way to identify where funds went, searchable by the final recipient, regardless of how many different budget items the outlay was recorded under. By linking these items to real-time execution data, the model enables observers to track actual outlays rather than merely the official presentation.

Tracking expenditures – including the full breakdown of allocations, final beneficiaries, and relevant execution details – is among the most prominent features of deep transparency. For instance, when a budget line item allocates public funds to a private entity for the provision of some public good (such as the construction of a new public hospital, road, school, stadium, sewage system, power grid, etc.), selecting the corresponding online entry would reveal a detailed breakdown of expenses for each component (e.g., construction costs, hospital furniture, medical equipment, interior fit-out, etc.). Our model could be applied with equal efficiency to off-budget items such as contingency funds, debt obligations, subsidies and grants, special accounts, tax breaks, and quasi-fiscal activities of state-owned enterprises (IBP, 2011).

To illustrate how this would work in practice, consider a hypothetical €10 million road construction contract awarded to a private firm. Under deep transparency, the budget system would record the commitment at contract signing, link it to the procurement record and the winning bid, trace each payment milestone to the contractor's bank details, and make the entire chain – from budget line to final recipient – permanently verifiable online in real time. No equivalent mechanism exists in current PFM or transparency frameworks.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

The practical relevance of such traceability is evident in a recent Hungarian case. Early research (Fazekas, King and Tóth, 2013) revealed – that Hungarian public procurement exhibited systemic, politically driven patterns in which contract allocation depended more on connections than on competition, with outcomes shifting alongside changes in political power – was later quantified by the Financial Times, which found that 14% (€28 bn) of all funds awarded in state tenders under Viktor Orbán went to 13 of his associates, based on an analysis of nearly 350,000 contracts (Caruana Galizia, 2026). Under a deep transparency model, patterns of this kind would become publicly visible within a few clicks.

Our model also enables object-based outlay aggregation, providing a quick, efficient, and straightforward method to calculate public expenditures by summing costs across various administrative units – such as the executive, courts, and regulatory institutions – based on specific activity categories. To illustrate, consider the U.S. federal budget, where outlays are categorized using object classifications, a system defined by the Office of Management and Budget. This system groups spending by transaction type, regardless of the agency incurring the expense, ensuring consistency across federal agencies and facilitating cost comparison. For instance, one can simply refer to the relevant object class across all agencies and departments to determine total expenditures on specific activities, such as travel (object class 21) or grants (object class 41).

This information is essential not only for budget transparency but also for responsible public spending. It lets citizens see how much different administrative or economic activities cost the government. While many national budgets provide access to this data, calculating aggregate sums for specific activities often requires tedious manual work. Our online model automates this process, allowing users, particularly MPs, to obtain these aggregates with just a few clicks.

Deep transparency enables interested parties (who may have different authorisation levels) to verify that all changes in the budget are permanently recorded using WORM technology, similar to tracking changes in a word processor, history on Wikipedia, or blockchain transactions. This refers to comprehensive information regarding the annual budget, ongoing revisions, and amendments, encompassing all aspects of budget item execution and reallocations. When a payment is made from the budget, the system traces the action, recording who made the payment, when, and why. If the budget is rebalanced – as it often occurs in the second half of the year – the changes are visible, allowing the public to judge which items were decreased or increased. Most importantly, if someone makes unauthorised changes or pays out more than the budget item allows, the trace remains, enabling oversight institutions (such as the audit office) to identify and hold the responsible parties accountable.

The model enables governments to disclose summaries of PPP finances, covering rights, obligations, and anticipated financial transactions. PPPs have become essential to fiscal transparency, with the 2019 IMF Fiscal Transparency Code requiring disclosure of financial details by private agents to ensure transparency and accountability (IMF, 2019: 28). Under this framework, governments must

PUBLIC SECTOR ECONOMICS
50 (3) 405–429 (2026)

disclose PPP financial summaries, including rights, obligations, and expected financial transactions, highlighting the fiscal impact without revealing sensitive, proprietary, or operational contract details.

Unlike typical public expenditures (such as the cost of public administration), which appear as immediate costs in government budgets, PPPs often conceal their financial impact by spreading costs over time or keeping them off the official balance sheet. These deals often create off-the-books obligations, such as guarantees for future losses or required payments when the government is the only buyer. In the end, this "private" debt can be just as significant as traditional public debt. To ensure transparency, all documents related to these PPP obligations should be stored in a public archive, in the same way as regular budget data.

The openness envisioned by deep transparency does not extend to personal data such as salaries or private information about public officials, nor to financial details tied to national security or sensitive business information of private firms involved in PPP programs. It does not require ministries or agencies to disclose information that is not already legally public. It works by integrating existing budget, procurement, and PPP contract data that are currently scattered across separate platforms and documents, making it difficult for non-insiders to see how specific outlays correspond to individual contracts. By linking these already-public records, the model makes execution traceable without expanding disclosure obligations or violating privacy and security rules. In line with the IMF Fiscal Transparency Code (2019), the government is not required to disclose proprietary, commercially sensitive, or operational details embedded in full PPP contracts. Such undisclosed elements may include proprietary technologies or processes, detailed financial models, specific operational procedures, pricing strategies, or security measures essential to project integrity. These exclusions are intended to protect private partners' competitive position, preserve intellectual property, and ensure the security and effectiveness of PPP arrangements.

That said, deep transparency requires governments to disclose information essential for public accountability without compromising commercial confidentiality. To ensure transparency when public funds are allocated to private agents for the provision of public goods, it is crucial to disclose the total contract value, the bids submitted by unsuccessful competitors, and lump-sum prices of key cost components. For example, if a private company is contracted to build a hospital, the budget should publicly list overall costs for major inputs such as construction work, labour, and core materials like concrete, bricks, or medical equipment. While individual salaries need not be published, aggregated input costs must be available so that concerned stakeholders can benchmark them against prevailing market prices.

Most importantly, deep transparency is crucial for identifying patterns in which a private firm repeatedly wins contracts, possibly due to political connections rather than competitive merit. This issue is especially notable in hybrid regimes and weak democracies, where firms often leverage political ties to secure contracts, but it is observable in developed Western democracies as well (Fazekas and Tóth, 2016).

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

3 POLICY INITIATIVES TOWARD TRANSPARENT BUDGETING

Over the past several decades, public financial management (PFM) frameworks have developed along two main lines. First, they establish standardised systems for governments to record, track, and manage public finances internally. These systems – such as treasury operations and accounting structures – are typically managed by Ministries of Finance and implemented through integrated platforms such as IFMIS, which connect budgeting, treasury, accounting, and reporting functions. They allow entrusted public officials to monitor budget execution across stages (commitment, accrual, payment) and ensure consistency, control, and auditability. International accounting standards such as IPSAS further underpin these systems by providing standardized frameworks for general-purpose financial reporting, improving the quality and comparability of public sector financial statements. However, this internal visibility is primarily designed for administrative and oversight purposes, not for direct public access. Only privileged public officials have the full picture of what was spent.

Second, these frameworks define what information should be disclosed externally. Governments are expected to publish structured reports, such as financial statements, budget execution reports, and fiscal summaries, that include data on revenues, expenditures, assets, liabilities, deficits, and fiscal risks. These practices are also reflected in IMF guidance on digital PFM systems, which emphasizes interoperability and standardised reporting across financial functions (IMF, 2019). Initiatives such as the Open Budget Index (OBI) and PEFA assessments build on these standards to evaluate the availability and quality of such disclosures. While these measures improve transparency, reporting remains aggregated, periodic, and institutionally mediated, without transaction-level traceability.

In practice, the implementation of these standards varies considerably across countries. Most advanced economies, particularly OECD countries, operate integrated PFM systems based on IFMIS platforms that generate detailed internal, transaction-level data and produce standardized fiscal reports for external use. However, even in these systems, public disclosure remains largely aggregated and periodic, with limited ability to trace specific expenditures to final beneficiaries. In many developing and transition economies – including the Western Balkans and Sub-Saharan Africa – PFM systems are not fully integrated, reporting practices are uneven, and transparency is often weaker or inconsistently applied. Despite these differences, a common feature across contexts is the gap between internally available financial data and publicly accessible data, particularly at the level of individual transactions.

Over the past two decades, a number of initiatives have sought to make public budgets easier for citizens to access and understand. National efforts such as Italy's Cohesion Policy portal, Barcelona's Budget Obert, and the USAspending.gov platform in the United States have made progress in publishing data on public expenditures. In Central and Eastern Europe, subnational budget watch initiatives

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

in countries such as Croatia, Macedonia, and Ukraine have similarly attempted to improve citizen access to local government finances, documenting persistent gaps between formal disclosure requirements and actual public access (Ott, 2006).

Among the most advanced PFM systems is Brazil's SIAFI (*Sistema Integrado de Administração Financeira*), an integrated treasury platform launched in 1987 that provides real-time tracking of federal budget execution across all ministries. SIAFI represents a significant achievement in internal fiscal control and is widely cited as a model for integrated financial management. Yet even SIAFI falls short of what deep transparency envisions: its transaction-level data serve internal government purposes and is not structured to allow citizens to independently verify payments to final beneficiaries. Like other advanced PFM systems, it improves the state's internal legibility of spending without creating public verifiability of execution. Yet across these efforts, the contribution remains largely limited to improving access. Real-time monitoring, systemic oversight, and accountability mechanisms often remain out of reach.

More systematic approaches have emerged from multilateral organizations and expert networks. Among these, the Open Budget Index (OBI), the International Monetary Fund's Fiscal Transparency Code, and the OECD's Budget Transparency Toolkit stand out. These initiatives have helped define fiscal transparency, but they also highlight persistent gaps in practice.

The OBI, developed by the International Budget Partnership, evaluates national budget transparency by assessing the availability and quality of eight core documents, ranging from pre-budget statements to audit reports. These evaluations follow a standardized methodology and include peer reviews, feedback loops with government authorities, and input from civil society researchers. The OBI's strength lies in its ability to normalize transparency as a measurable standard. Recent cross-country evidence using OBI data suggests that higher levels of budget transparency are positively associated with financial sustainability outcomes, including fiscal balance and debt performance. However, the model is largely document-based and episodic. Governments may receive high scores for publishing planned budgets, while withholding information on mid-year expenditures or execution reports. More than 40 percent of countries assessed by the OBI fail to release such updates.

Disclosure of PPPs and off-budget liabilities is even more limited. As a result, the OBI reveals what governments are willing to disclose at selected points in time, but not how public funds are managed or disbursed throughout the fiscal year. Notably, the IBP's own companion guide explicitly identifies five domains as most vulnerable to opacity and abuse: extra-budgetary funds, tax expenditures, quasi-fiscal activities, contingent liabilities, and future liabilities (IBP, 2011). These are precisely the domains that constitute the shadow budget as defined in this paper. Deep transparency is designed to make these domains traceable and verifiable in real time – going beyond what the OBI framework can achieve through periodic document-based assessment alone.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

The IMF's Fiscal Transparency Code, last revised in 2019, offers a broader framework organized around four pillars: fiscal reporting; forecasting and budgeting; risk analysis and management; and resource revenue oversight. It encourages the publication of aggregate fiscal data, including contingent liabilities such as those stemming from PPPs. Its emphasis is on macroeconomic stability and improved governance, achieved through timely and credible fiscal reporting. Yet the IMF Code remains centred on aggregate data and retrospective assessments. It does not require mechanisms that would allow transaction-level visibility or real-time oversight by the public or independent auditors. The primary objective is to make fiscal risks visible to markets and oversight institutions, rather than to enhance day-to-day accountability in public spending.

The OECD's Budget Transparency Toolkit (2017) takes a more operational approach. It emphasises not only openness but also integrity and accountability. The Toolkit recommends publishing detailed information on budgets, contracts, procurement procedures, and performance metrics. It also addresses the specific risks posed by opaque financial arrangements, especially those involving PPPs and similar off-budget commitments. Still, the Toolkit remains a set of best practices. Adoption is voluntary, enforcement mechanisms are absent, and data availability remains uneven and often delayed. Like the OBI and the IMF Code, it is anchored in what governments choose to disclose, not in mechanisms that ensure continuous and verifiable transparency.

These initiatives have succeeded in setting global standards and in defining fiscal transparency as a policy objective. Yet they share two critical shortcomings: they focus on selected publication rather than on implementation, and they depend on periodic reporting rather than on continuous observation. They describe what should be public but offer little insight into whether those principles are honoured in practice.

Deep transparency differs from all of these frameworks in a fundamental way. The OBI, IMF Code, OECD Toolkit, and PEFA assessments are, at their core, disclosure diagnostics: they measure and encourage what governments choose to publish, in the form of documents, reports, and aggregated data. They are institutionally mediated – citizens receive information filtered through government reporting choices, released periodically, and rarely traceable to individual transactions or final recipients. Public procurement and PPPs, which now account for up to 33 percent of global GDP (Fazekas et al., 2024), remain among the least transparent domains even under these frameworks. Deep transparency is something different. It is a democratic tool. It does not ask governments to improve reporting – it makes the underlying data directly and permanently verifiable by any citizen, civil society organization, or oversight body. Where existing frameworks produce document-level transparency, deep transparency produces execution-level traceability. Where existing frameworks rely on government willingness to disclose, deep transparency creates an architecture in which the link between each budgetary outlay and its final private recipient – the vendor, contractor, or beneficiary who ultimately receives public

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

funds – is online, immutable, and open to independent verification. This is the shift from publication to verification, from periodic reporting to continuous accountability, and from institutional mediation to direct democratic scrutiny.

Addressing these limitations requires a shift from standard-setting to infrastructure-building. A model that links budget allocations directly to public contracts and expenditures, stored in tamper-resistant systems such as WORM or sealed by blockchain technologies, allows for real-time, transaction-level traceability. This model does not replace the OBI, IMF Code, or OECD Toolkit. It puts their principles into practice by enabling real-time, traceable spending so that citizens can, on this basis, make informed political and policy decisions.

4 A TECHNICAL SOLUTION FOR DEEP BUDGET TRANSPARENCY

Multiple datasets define the budget and the recipients of its money. Some are linked to the government portal, while others are stored by various governmental agencies or local authorities. Gathering all corresponding budgeting data can ensure transparent and efficient access for all interested parties. The qualified majority of the parliamentary Security and Defence Committee should identify datasets that cannot be made public. Those data should be available to authorised persons, such as MPs, under appropriate procedures.

It bears repeating that the proposed architecture applies uniformly to all categories of discretionary and off-budget spending, not only to public procurement or PPP-related items. Any financial outlay – whether a subsidy, grant, transfer, routine operational payment, or contractual obligation – enters the system as a uniquely identified budget item with a timestamp and metadata. This guarantees that subsidies, special-purpose transfers, and administrative reallocations are captured by the same version-tracking logic as procurement transactions. Every modification, reallocation, or release of funds is recorded through the interoperability layer and written once to the WORM-secured Master node, ensuring a verifiable audit trail across the entire execution phase of the budget.

The new solution should enable monitoring of the complete budget cycle by implementing advanced search and data analytics. To accomplish this, we propose introducing an integration layer for existing datasets and services and suggest the application of several innovative technologies to improve problem-solving and resource utilisation. In addition, we intend to accelerate the digitisation process of public administration as a preferred means to hold the government accountable.

As the first task, we ensure interoperability between different participants in the budget creation process. Therefore, our solution (figure 1), based on Interconnected government services (Bojović et al., 2023), incorporates three core principles:

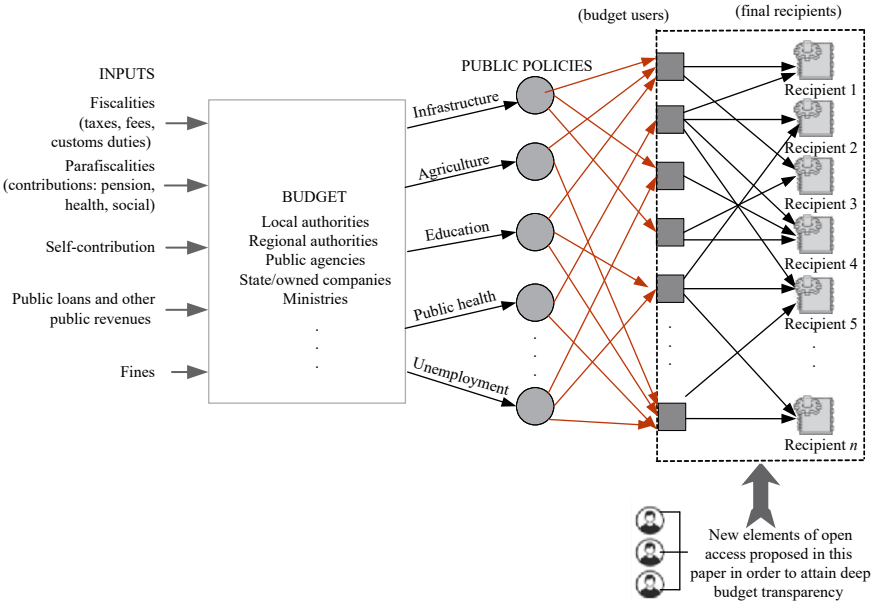
– Using the available computing infrastructure without investing in new storage capacity (data remain in their current locations, entirely functional and self-sustaining) or spending time on data preparation.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

– Decoupling data processing from data storage. This enables an application to communicate with a single instance despite multiple databases (assigned to different budget entities such as local, state, public agencies, and others) reading the data. Each user represents a single instance that can access the same application using their identification parameters.
– Dynamic Budget Reporting: This approach facilitates the retrospective review of all current public budget execution versions, providing insight into historical modifications. It enables authorities within the budget's jurisdiction to decide on project financing. The analysis evaluates whether contract prices align with market rates, aiming to prevent inflated costs and ensure compliance with legal requirements.

FIGURE 1

Public budget allocation process with the final beneficiaries



Source: Authors.

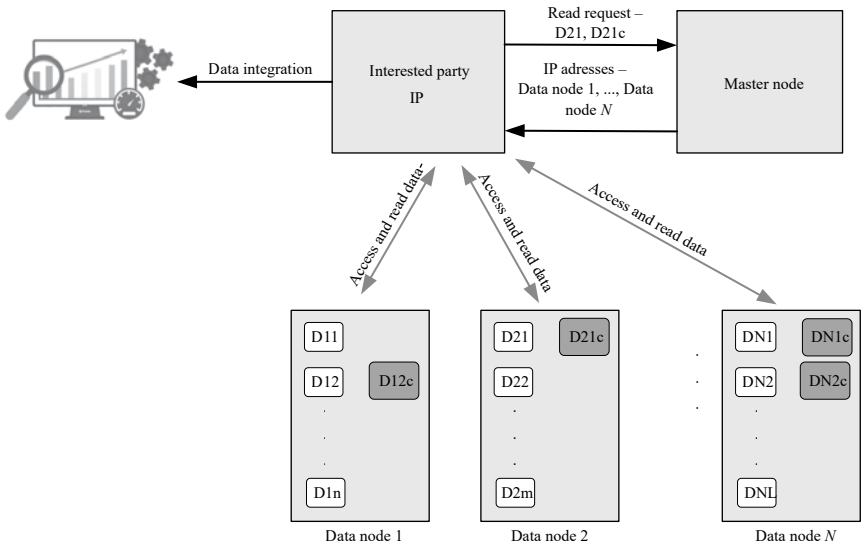
As a common access point, we offer an open data portal containing information on the annual budget, current changes, and amendments (all execution versions of budget items and positions that initiate redistribution or engagement of new funds). Experience has shown frequent changes in the government's priorities during the mandate. These changes usually result in new policies for allocating and spending budget funds (e.g., additional funds for social assistance, unemployment, eliminating the consequences of a catastrophic event, and spending on pension or defence). For this reason, we propose integrating the solution shown in figure 1 into the e-government system. By digitising budgetary procedures, we can realise the critical objectives of good governance.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

As depicted in figure 2, our solution implies building a distributed file system with a Master/Slave architecture comprising a single Master node and several Slaves (so-called Data nodes). As a multifunctional and highly available server, the Master node is responsible for maintaining and managing data on the Data nodes. Access to data is possible only after obtaining permission from the Master node. It contains the Namespace entity, which acts as a container, with file name grouping, metadata, and information such as file owners, access permissions, data locations, and data size. We must emphasise that the actual budget data resides on Data nodes, while a separate block called budgeting – historical data stores previous execution versions and budgeting rules.

FIGURE 2

A technical solution for deep budget transparency



Source: Authors.

The state budget allocates funds to budget users (BU1, BU2, etc.) (see figure 1). These budget users, in turn, distribute the allocated funds to specific entities responsible for implementing policy programs (D11, D12, ..., DNL). The Master node, a hardware system designed to log every transaction, records all financial transactions, even those that might be altered or erased locally. When introducing a new version of a budgetary item, such as D12, following a public procurement process, it requires reporting to the Namespace Data Management unit (NDM). This report includes essential data such as the amount paid and the recipient, along with relevant metadata: the initiator of the transaction, the precise transaction time, and any restrictions on data access (these restrictions are essential for state security reasons and must be pre-approved). NDM then registers this data and metadata into the Master node. The newly corrected version, D12c, replaces the original D12 locally, but the Master node permanently stores both versions.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

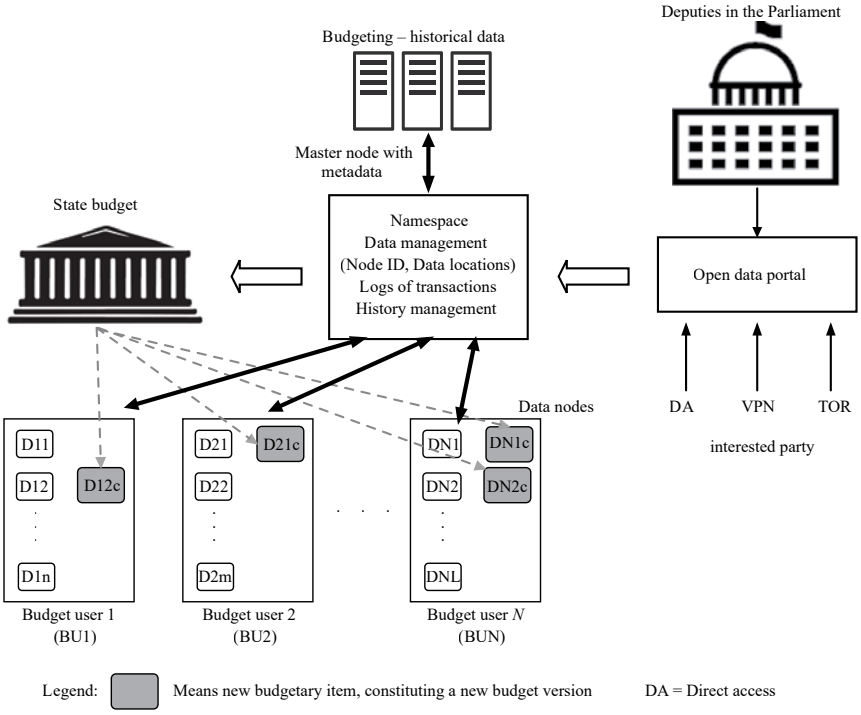
Our solution centres on maintaining data integrity in the Master node using WORM technology. This technology involves writing data to a storage medium only once. Once in the WORM database, the data becomes immutable and permanent, preventing any possibility of erasure, modification, substitution, or re-encryption. The previous and current data remain recorded in the local database, while the Master node documents every change to the metadata files. The primary goal of this proposed solution is to ensure that any alteration, at any level, remains permanently visible, thereby guaranteeing full transparency in the use of budget funds and enhancing accountability for government institutions. Furthermore, data integrity can be further secured, through the optional use of blockchain technology, which can also be used as an alternative to WORM storage. Though in the implementation of the Deep Transparency model one can opt for either WORM or blockchain, it would be beneficial if they could be used together, given that the passage of time always brings new possibilities of cyber attack, and that redundancy in security costs little to implement and could save a lot in the unforeseen future.

The parliament icon in the upper-right corner of figure 2 signifies that data is transparent to MPs through the secure parliamentary portal. It is also accessible to any interested party (IP) based on their access rights (similar to a Role-Based Access Control system), which may be more restricted than those of MPs for national security reasons. IPs can access the Open Data Portal directly or, if they prefer anonymity, through a Virtual Private Network (VPN) from any public space, such as a café, or even via TOR (The Onion Router) for complete anonymity and security. Implementing a VPN before accessing TOR is also advisable in such cases because a VPN prioritises data security and cannot fully anonymise IP identity. Using VPN over TOR makes it extremely difficult to trace the origin of the request and the data back to the user (the original IP address remains hidden).

Figure 3, which presents a more technically demanding aspect, indicates that our search engine should be capable of providing this functionality according to the specified procedure:

- the MP or any authorised person (IP, who, in most cases, can be anyone) can access Master node, asking for the metadata related to the desired data
- the Master node will return the lists of data nodes with the data locations
- after that, IP connects to data nodes, where the data are stored
- the IP can start reading data in parallel from all corresponding data nodes
- once the IP gets all the required data, it can combine them into a report.

FIGURE 3
Data search engine



Source: Authors.

The first step in this procedure is essential for full transparency of the proposed solution. It implies clearly defined data access rights within the system. Next, IP and budget creators must undergo authentication (the process of verifying a user's identity) and authorisation (a process giving official permission). Finally, the system grants specific permissions based on data ownership, group membership, or to anyone for data designated as fully open (the majority of data should fall in the latter category).

Use case 1: Budget creators at different levels (ministries, local authorities, state-owned companies, and public agencies) own their data. They enter data into the appropriate database without deleting or modifying it, in accordance with the principles of WORM technology. The Data node will record every change as a new execution version. The Master node database will also automatically record metadata that more precisely identifies the change (what, when, and who made it). Every outlay from the budget needs to be associated with a budget item. If none exists, the database operator can create one, with a remark explaining who, why, when, and according to which document authorised it.

Use case 2: IPs can access data after the authentication and authorisation procedure. They can read data according to the IP's permission level and search according to criteria. The proposed solution enables defining rights at both individual and group levels – group permissions are used for confidential data.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

It is necessary to emphasise that the Master node plays a crucial role in our solution. It contains a daemon (a software component that runs processes in the background of the Master node) responsible for maintaining and managing the data nodes and metadata stored in the Master node files (such as the location of data, size of data, permissions, and hierarchy). Two files are associated with the metadata:

– File system state: contains the entire state of the Namespace (from the creation of the Master node onward), with clearly indicated checkpoints where data changes occurred.
– Edit Logs: records each change to the metadata (checkpoints), e.g., a deletion of an item in the file system. The Master node will immediately record this in the Edit Logs.

To improve the database's resistance to unauthorised changes, we propose introducing a blockchain that protects the data hash and metadata of all previously introduced records so that anyone who reads the data can verify the consistency of all previous data.

We differentiate among at least five authorisation levels in the Namespace directory:

1) Read – citizens,
2) Read – authorised persons (for restricted data, such as military spending, recipients of social subsidies, and similar),
3) Read and comment (MPs; authorized civil society agents),
4) Read, comment, and challenge (official oversight institutions; public prosecution),
5) Read, comment, and write (budget officials, policymakers, top public servants).

Each authorisation level is a file containing the list of users with applied rights and passwords for some budgetary items.

From an implementation perspective, deep transparency relies on three interconnected cost components: infrastructure, data integration, and institutional adaptation. A fundamental requirement is establishing data integrity and immutability layer, centred on hardware-level WORM storage, with optional blockchain-type distributed ledgers (€100,000 to €200,000). The initial investment for this component typically ranges from €150,000 to €350,000, laying a technical foundation for public trust. Upgrades to legacy IFMIS platforms are essential to enable real-time, API-based interoperability. The costs associated with these upgrades, including data warehousing and portal development, are estimated to be between €1.2 million and €2 million. An additional €500,000 to €800,000 will be required for data cleaning and alignment with international standards, such as the Open Fiscal Data Package (OFDP). Institutionally, the Ministry of Finance must serve as the central anchor, playing a crucial role in enforcing data-sharing mandates and ensuring inter-agency coordination. Thus, capacity building and change management are critical non-technical cost drivers, estimated at €400,000 to €700,000.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

Annual operating costs, covering cybersecurity and continuous data validation, range between €300,000 and €500,000. To mitigate implementation risks, we recommend a phased pilot program in high-expenditure sectors, such as health or infrastructure. Even under worst-case assumptions, total lifecycle costs over five years are projected to remain below €5 million – typically lower than the fiscal losses resulting from a single large-scale corrupt procurement.

5 POLITICAL CONSEQUENCES OF DEEP TRANSPARENCY

Implementing deep transparency changes the political economy of the shadow budget and discretionary spending. By recording each outlay as it occurs, the model reduces the discretionary space through which political actors distribute favours, reward loyalty, or hide losses. Informational power shifts: what was previously known only within patronage networks becomes visible to oversight bodies and, in filtered form, to the public. These shifts impose political costs on actors who rely on opacity, and raise the stakes for institutions expected to enforce disclosure. Since reducing transparency, clientelism, and discretionary abuse presupposes that budget data become accessible to far more actors – including ordinary citizens – we first address a preliminary question: whether such expanded visibility risks overwhelming users through technical complexity and information overload.

These concerns are substantial. Technical complexity (Heald, 2015) and information overload (Heald, 2006) pose real obstacles to expanding public access to budget data. Deep transparency overcomes these challenges by filtering information for different audiences and clarifying budget details. This is crucial, as excessive or unstructured citizen participation – when paired with unfiltered transparency – can undermine fiscal performance and lead to short-term populist pressures (Jung and Kim, 2025). Yet at the opposite end of the spectrum, the absence of citizen involvement can be equally harmful. The COVID-19 pandemic revealed how emergency spending can bypass conventional transparency mechanisms, leading to large off-budget expenditures, minimal legislative or audit oversight, and almost no public participation in fiscal decisions – despite the unprecedented scale and urgency.

Deep transparency is designed to avoid both of these pitfalls. It emphasises institutional design – clarifying, structuring, and tailoring data to ensure that transparency serves accountability without descending into either technocratic opacity or populist distortion. We avoid the transparency illusion trap (Heald, 2003), in which increased transparency is only effective with proper data interpretation and dissemination mechanisms. Our model enhances external accountability by releasing data in a reusable format, allowing citizens, NGOs, and vulnerable groups to engage with budgeting and public policy (Anessi-Pessina et al., 2016). Knowledgeable citizens and organizations act as "information brokers," simplifying and communicating complex information. Additionally, recent AI developments enable summarisation of budget items by filtered parameters, assisting users in understanding detailed information. In short, technological advancements allow our transparency model to help citizens better understand budgets.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

The greatest gains from deep transparency arise where discretion is broad and oversight weak: in the political economy of off-budget spending and inside public administration. Once every budgetary move becomes visible and traceable, incentives shift, professional integrity matters more, and the patterns of patronage that thrive in the dark begin to lose ground.

Some empirical research has already demonstrated that civil service personnel's professional integrity substantially impacts budget allocation and wasteful spending. A study by Bostashvili and Ujhelyi (2019) found that in electoral years, high party patronage (indicative of unprofessional civil servants) led to a 12 per cent increase in wasteful spending in the electoral year, while in the pre-election years, the increase was 9 per cent. Conversely, public officials who demonstrate professionalism (low party patronage) are less likely to prioritize political interests by manipulating government spending to align with electoral needs and vested interests. In this context, deep transparency empowers these public officials to avoid the misuse of public funds, as the cost of such misconduct increases exponentially. When public officials are equipped with a deep transparency budget tool that entails tracking all budget changes (section 4), their ability to serve political interests plummets.

Besides reducing budgetary spending to enhance the incumbent government's chances in the impending elections, we expect deep transparency to function not only as a constraint on corruption but also as a barrier to entry for those inclined toward it. By rendering abuse traceable, costly, and institutionally unforgiving, it alters the calculus of opportunists. The logic is simple: when the space for discretion shrinks and the likelihood of exposure rises, those who rely on opacity will seek other avenues (UNODC, 2020). Recent political economy research suggests that institutional environments characterized by widespread corruption tend to attract individuals with a higher propensity for dishonest and self-serving behaviour.

Experimental studies from India reveal that students aspiring to civil service positions – operating within a high-corruption context – exhibit lower levels of altruism and a greater willingness to cheat or engage in bribery, consistent with a self-enrichment motive (Banerjee, Baul and Rosenblat, 2015; Hanna and Wang, 2017). Moreover, a corruptive environment can change behaviour even among individuals who enter public administration without corruptive preferences. This is documented by Gans-Morse et al. (2020), who show that corruption in Russia arises not from who enters public service, but from how institutions transform individuals once they are inside.

A similar claim – that a non-corruptive environment will make political parties look for people who have non-corruptive preferences – applies to political parties. Office-seeking parties, which aim to win elections and implement policies, tend to attract highly educated, experienced, and ideologically oriented activists who play influential roles in the party (Scarrow, 2015). In contrast, parties operating under a corrupt system do not actively seek out such members. Research indicates that

high levels of clientelism can significantly impact party organisation, including its membership structure, leadership, and inner-party democracy (Chiru, 2024). In extreme cases, these parties may collaborate with actors who provide violent services, such as criminal groups, assassins, enforcers, private security companies, thugs, political protesters, football hooligans, and paramilitaries (Pavlović, 2022; Tomić and Pavlović, 2023).

While deep transparency as a full system has not yet been empirically tested – which is precisely why this paper is written – evidence from adjacent domains offers useful proxies for the plausibility of these effects. A first proxy comes from research on self-selection into public institutions: studies conducted in Denmark show that students aspiring to public-sector employment demonstrate greater honesty and pro-social preferences in low-corruption environments, suggesting that clean institutions select for clean agents (Barfort et al., 2019). This supports the expectation that reducing discretionary space through deep transparency may, over time, alter who is attracted to public office.

A second proxy comes from the domain of party finance transparency. Casal Bértoa, Heapy-Silander and Lynge (2025) show that financial disclosure requirements, combined with independent oversight, reduce political corruption and increase party institutionalisation – as evidenced by the Latvian case, where the establishment of the Corruption Prevention and Combating Bureau (KNAB) as an independent oversight body, combined with mandatory party finance disclosure requirements, produced a measurable reduction in political corruption and a sustained increase in party institutionalisation over two decades. If transparency in party finances alone can alter party behaviour and incentive structures, the more comprehensive and tamper-resistant architecture of deep transparency may plausibly be expected to produce analogous effects at the level of budget execution. Taken together, these proxies are indicative rather than definitive. They suggest that deep transparency is expected to reduce discretionary space, which in turn may shift incentives, contributing to lower corruption and improved democratic outcomes. Deep transparency can therefore plausibly change politicians' motives, transforming them into more responsible officials. By offering a tracing mechanism for executing each budget item, including details on contracting out public services by linking budget items to public contracts, deep transparency can alter the incentives for politicians to engage in corruption and state capture.

We now turn to the question of our model's applicability. Is deep transparency more suitable for developed economies and democratic political systems, or for undeveloped and non-democratic systems where the cost of office abuse is low? Previous research (Khagram, Fung and de Renzio, 2013) shows that greater fiscal transparency does not result solely from institutional design. It is typically precipitated by political ruptures – transitions from authoritarian rule, fiscal and economic crises, corruption scandals, or international pressures that empower domestic reformers. These events alter the distribution of power, creating openings for actors who

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

demand access to budgetary information. The historical record supports this pattern. The post-communist transitions in Central and Eastern Europe after 1989 produced sweeping budget transparency reforms alongside new constitutions and freedom-of-information laws. The IMF Fiscal Transparency Code itself – introduced in 1998 – was a direct response to the Asian financial crisis of 1997, which exposed the dangers of opaque public finances and created international pressure for disclosure norms. Brazil's SIAFI system, one of the most advanced integrated budget execution platforms in the world, was introduced in 1987 as part of the democratic transition following military rule. In each case, rupture preceded reform. We expect the same logic to apply to deep transparency: it is most likely to emerge not through gradual institutional evolution but in the wake of fiscal crises, anti-corruption breakthroughs, or democratic transitions that shift the balance of power decisively toward reformers demanding verifiable public accounts.

Having considered where deep transparency may be most feasible – though politically obstructed – we now turn to a different question: do more developed democracies, where implementation faces fewer barriers, also have a pressing need for it? Our model is much easier to implement in developed economies and democratic systems. Although they are generally more open and less corrupt, they could still benefit from deep transparency. Incumbent governments use public funds for overt or covert campaign spending in election years everywhere. Developed economies often exhibit subtler forms of corruption, embedded within formal institutional frameworks rather than through overt violations of law. Two prominent mechanisms are regulatory capture and revolving door employment (Tomić, 2023; Ang, 2024). The former enables industries to shape regulations to their advantage, while the latter involves public officials moving between government and private sector roles, using privileged access to favour connected firms. These practices significantly affect budgetary decisions – particularly in areas like public procurement and PPPs – leading to inefficiencies and misallocation of public funds. Moreover, recent empirical evidence supports this claim: the new state capture index (Kaufmann, 2024) shows that abuse of public resources is rising even in advanced democracies and developed market economies. While state capture was once considered a post-communist phenomenon (World Bank, 2002), current data suggest it has become a global trend, indicating that all systems – regardless of income level or regime type – stand to gain from adopting deep transparency practices.

6 CONCLUSIONS AND FUTURE RESEARCH

This article has argued that the weakness of contemporary budget transparency is not only that governments disclose too little, but also that what they disclose is typically aggregated, delayed, and difficult to verify. Our contribution is to propose a model of deep transparency that addresses both problems. It makes public spending more visible by linking budget allocations to final recipients, including private actors, and it makes budget execution traceable in real time through tamper-resistant digital infrastructure. It also makes independent scrutiny possible: citizens no longer need to wait for governments, auditors, or intermediaries to tell them how public money was spent. They can examine the record themselves and draw their own conclusions.

PUBLIC SECTOR ECONOMICS
50 (3) 405–429 (2026)

In the concluding remarks, we address several remaining issues for future research, the most important of which are transparency on the revenue side and the appropriate extent and type of budgetary details that should be disclosed.

Deep transparency focuses exclusively on the expenditure side of the budget, leaving the revenue side outside its current scope. In line with the IMF Fiscal Transparency Code (2019), which calls for timely and detailed disclosure of revenue sources, tax expenditures, and natural resource income, we acknowledge that transparency in revenue collection is equally essential for fiscal accountability. Future research should extend this framework to include mechanisms for tracking tax policy impacts, public reporting of tax breaks, and the governance of non-tax revenues – especially those most vulnerable to misuse, such as off-budget and natural resource revenues.

As mentioned, recent policy and academic research trends in budget transparency indicate a movement towards more open budgets, with increasing details disclosed to the electorate and interested parties. Despite substantial efforts to implement these concepts, achieving a satisfactory level of transparency in budgeting, including online transparency, remains challenging (Rubin, 2020). Deep transparency implies a significant level of detail regarding PPPs and budgetary changes that can be monitored in real time. Is this level of detail too much?

The usual argument is that excessive budget openness can present significant challenges and be counterproductive (Rubin, 2020). Overloading the public with disaggregated data can obscure rather than clarify fiscal information, making it difficult to interpret. Too much information can result in public and interest-group demands that hinder effective decision-making, such as reducing spending or eliminating tax breaks. Additionally, open processes can prolong decision-making and increase contention, raising unrealistic public expectations and causing widespread frustration. Therefore, balancing transparency with practicality is essential to maintain effective governance and public trust.

We argued that rapid technological advancement would enable our deep transparency model to filter and summarise the significance of each disclosed item. For example, with one click, it would be possible to obtain information on how much public money a specific private entity received in a particular year, or how much was spent overall on asphalt, concrete, disinfectants, or printer paper.

However, we did not address which items should fall under the principle of openness. Every treasury contains many off-budget items, such as tax breaks and state-owned enterprises (SOE) spending (IBP, 2011). Tax breaks are among the most effective off-budget means of preferential treatment for social, economic, and political interests, as well as of budgetary abuse. Because they are less visible than direct spending, they can obscure the actual financial state of a budget. Similarly, in countries where SOEs provide many public services, including their spending in budget transparency is essential.

PUBLIC SECTOR ECONOMICS
50 (3) 405-429 (2026)

Deep transparency enables access to this information, but the format of such a budget must be conceptualised from the start. Our model can link existing budgetary items with off-budget items. For example, if a government takes a loan from an international financial institution to cover an SOE's losses, the loan will be linked to the SOE's financial statement in which the loss appears. It is difficult to say more than this at present.

Finally, will deep transparency, even if implemented, necessarily lead to more accountability? The evidence from previous advancement in budget transparency exists but remains limited (Khagram, Fung and de Renzio, 2013). Isolated successes – such as Uganda's reduction in school grant leakages or Brazil's electoral penalties for corrupt officials – are exceptions rather than the rule. For transparency to have real impact, a full action cycle must be in place: data must be accessible and intelligible; actors must have incentives to respond; and other institutions must be capable of acting on their demands. Without this alignment, deep transparency becomes an empty gesture – form without consequence. Effective anti-corruption measures require more than simply making information available; they depend on robust conditions for publicity and accountability, such as education, media circulation, and free elections (Lindstedt and Naurin, 2010). Addressing these issues in the context of budget transparency requires a separate paper.

The shift from publication to verification is not a technical refinement – it is a democratic one. Governments already hold the data. The infrastructure to make it permanently and publicly verifiable exists. What has been missing is the institutional commitment to deploy it. That commitment matters now more than ever: state capture is rising across regime types, discretionary spending remains the principal vehicle for political abuse, and existing transparency frameworks have reached the limits of what disclosure alone can achieve. Deep transparency does not guarantee accountability. But it removes the condition on which unaccountability depends – the asymmetry between what governments know and what citizens can see.

Disclosure statement

The authors have no conflicts of interest to declare.

Declaration of generative AI and AIassisted technologies in the writing process

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

REFERENCES

1. Anessi-Pessina, E. [et al.], 2016. Public sector budgeting: a European review of accounting and public management journals. *Accounting, Auditing & Accountability Journal*, 29(3), pp. 491-519. https://doi.org/10.1108/AAAJ-11-2013-1532
2. Ang, Y. Y., 2024. Adaptive political economy: toward a new paradigm. *World Politics*, 75(5), pp. 1-18. https://doi.org/10.1353/wp.2025.a954433
3. Banerjee, R., Baul, T. and Rosenblat, T., 2015. On self-selection of the corrupt into the public sector. *Economics Letters*, 127, pp. 43-46. https://doi.org/10.1016/j.econlet.2014.12.020
4. Barfort, S. [et al.], 2019. Sustaining honesty in public service: the role of selection. *American Economic Journal: Economic Policy*, 11(4), pp. 96-123, https://doi.org/10.1257/pol.20170688
5. Bauhr, M. [et al.], 2020. Lights on the shadows of public procurement: transparency as an antidote to corruption. *Governance*, 33(3), pp. 495-523. https://doi.org/10.1111/gove.12432
6. Bojović, Ž. [et al.], 2023. Interconnected government services: an approach toward smart government. *Applied Sciences*, 13(2). https://doi.org/10.3390/app13021062
7. Bostashvili, D. and Ujhelyi, G., 2019. Political budget cycles and the civil service: evidence from highway spending in US states. *Journal of Public Economics*, 175, pp. 17-28. https://doi.org/10.1016/j.jpubeco.2019.04.004
8. Caruana Galizia, P., 2026. How Viktor Orbán's oligarchs reaped billions in public contracts. *Financial Times* (Accessed March 24, 2026).
9. Casal Bértoa, F., Heapy-Silander, A. and Lynge, H., 2025. Can financial transparency help parties institutionalize and save democracy? *Policy Studies*, 47(4), pp. 572-594. https://doi.org/10.1080/01442872.2024.2435380
10. Chiru, M., 2024. Clientelism, party organization and intra-party democracy. *Comparative Political Studies*, 58(4), pp. 680-713. https://doi.org/10.1177/00104140241252082
11. Curry, W. S., 2022. *Contracting for Services in State and Local Government Agencies: Best Practices for Public Procurement*. 3rd ed. New York: Routledge. https://doi.org/10.4324/9781003305910
12. Dávid-Barrett, E. and Fazekas, M., 2020. Grand corruption and government change: an analysis of partisan favoritism in public procurement. *European Journal of Criminal Policy and Research*, 26, pp. 411-430. https://doi.org/10.1007/s10610-019-09416-4
13. Fazekas, M. [et al.], 2024. Global contract-level public procurement dataset. *Data in Brief*, 54, 110412. https://doi.org/10.1016/j.dib.2024.110412
14. Fazekas, M. and Tóth, I. J., 2016. From corruption to state capture: a new analytical framework with empirical applications from Hungary. *Political Research Quarterly*, 69(2), pp. 320-334.

15. Fazekas, M., King, L. P. and Tóth, I. J., 2013. Hidden depths: The case of Hungary. *ERCAS Working Paper,* No. 36.
16. Gans-Morse, J. [et al.], 2020. Self-selection into public service when corruption is widespread: the anomalous Russian case. *Comparative Political Studies*, 53(2), pp. 234-270. https://doi.org/10.1177/0010414020957669
17. Hanna, R. and Wang, S.-Y., 2017. Dishonesty and selection into public service: evidence from India. *American Economic Journal: Economic Policy*, 9(3), pp. 262-290. https://doi.org/10.1257/pol.20150029
18. Heald, D., 2003. Fiscal transparency: concepts, measurement and UK practice. *Public Administration*, 81(4), pp. 723-759. https://doi.org/10.1111/j.0033-3298.2003.00369.x
19. Heald, D., 2006. Varieties of transparency. In: C. Hood and D. Heald, eds. *Transparency: The Key to Better Governance?* Oxford: British Academy, pp. 24-43. https://doi.org/10.5871/bacad/9780197263839.003.0002
20. Heald, D., 2012. Why Is Transparency About Public Expenditure So Elusive? *International Review of Administrative Sciences*, 78(1), pp. 30-49. https://doi.org/10.1177/0020852311429931
21. Heald, D., 2015. *Surmounting Obstacles to Fiscal Transparency*. Washington, DC: GIFT.
22. Hood, C., 2006. Transparency in historical perspective. In: C. Hood and D. Heald, eds. *Transparency: The Key to Better Governance?* Oxford: British Academy, pp. 2-23. https://doi.org/10.5871/bacad/9780197263839.003.0001
23. IBP, 2011. *Looking Beyond the Core Budget: Guide to Transparency in Public Finances.* Washington: International Budget Partnership.
24. IMF, 2019. *IMF Fiscal Transparency Code*. Washington: International Monetary Fund.
25. IMF, 2023. Digital Solutions Guidelines for Public Financial Management. *IMF Technical Notes and Manuals,* No. 2023/007. https://doi.org/10.5089/9798400251566.005
26. Jung, H. and Kim, B. H., 2025. Budget transparency and fiscal soundness: citizens' political participation as a moderating factor. *Public Money & Management*, 45(7), pp. 799-808, https://doi.org/10.1080/09540962.2025.2456123
27. Kaufmann, D., 2024. State capture matters: considerations and empirics toward a worldwide measure. In: S. Rose-Ackerman, ed. *Public Sector Performance, Corruption, and State Capture in a Globalized World*. Routledge: Oxon, pp. 207-237.
28. Khagram, S., Fung, A. and de Renzio, P. (eds), 2013. *Open Budgets: The Political Economy of Transparency, Participation, and Accountability*. Washington, DC: Brookings Institution Press.
29. Lindstedt, C. and Naurin, D., 2010. Transparency is not enough: making transparency effective in reducing corruption. *International Political Science Review*, 31(3), pp. 301-322. https://doi.org/10.1177/0192512110377602

30. OECD, 2017. *Budget Transparency Toolkit: Practical Steps for Supporting Openness, Integrity and Accountability in Public Financial Management.* Paris: OECD.
31. Ott, K. (ed.), 2006. *Making Public Finance Public: Subnational Budget Watch in Croatia, Macedonia, and Ukraine.* Budapest: Local Government and Public Service Reform Initiative, Open Society Institute.
32. Pavlović, D., 2022. How to approach state capture in post-communist Europe: a new research agenda. *Journal of Contemporary European Studies*, 31(3), pp. 960-978. https://doi.org/10.1080/14782804.2022.2106951
33. Rubin, I. S., 2020. *The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing.* 9th ed. Washington, DC: CQ Press.
34. Scarrow, S., 2015. *Beyond Party Members: Changing Approaches to Partisan Mobilisation.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199661862.001.0001
35. Tomić, S. and Pavlović, D., 2023. Blatant, not latent: patronage in top-level appointments in Serbia. *NISPAcee Journal of Public Administration and Policy*, 16(1), pp. 181-209. https://doi.org/10.2478/nispa-2023-0008
36. Tomić, S., 2023. Revolving door. In: R. Barrington [et al.], eds. *Dictionary of Corruption*. London: Agenda Publishing.
37. UNODC, 2020. Detecting and investigating corruption (Module 6). In: *UNODC Module Series on Anti-Corruption.* Vienna: United Nations Office on Drugs and Crime.
38. Wildavsky, A., 1964. *The Politics of the Budgetary Process.* Boston, MA: Little, Brown and Company.
39. World Bank, 2002. *Transition: The First Ten Years. Analysis and Lessons for Eastern Europe and the Former Soviet Union.* Washington, DC: World Bank.

Applying VAT to loan-deposit intermediation

GUILLERMO PEÑA, Ph.D.*

Article**
JEL: H25, H21, D61
https://doi.org/10.3326/pse.50.3.5

* The author expresses his gratitude for the funding received from the project entitled "Municipios y servicios financieros: shocks, ratios y efectos", by the Fundación Ibercaja/Universidad de Zaragoza (JIUZ2022-CSJ-19); the "S23_20R: Public Economics Research Group", "S39_23R: ADETRE", "PROY_S_15_24" and "LMP92_21: Estrategias frente a la despoblación: enfoque multidisciplinar para una gestión integrada, inclusiva y dinamizadora", both funded by the Government of Aragon, and the research project "CISIO": PID2024-157255NB-I00, by the Spanish Ministry of Science, Innovation and Universities. The author would also like to thank two anonymous referees for their comments and suggestions.

** Received: February 26, 2026
Accepted: June 29, 2026

Guillermo PEÑA
Department of Economic Analysis and IEDIS, Universidad de Zaragoza, Pedro Cerbuna, 12, 50009 Zaragoza, Spain
e-mail: gpena@unizar.es
ORCiD: 0000-0002-8816-5816

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

Abstract

Technical difficulties associated with transaction-by-transaction taxation methods have led financial services to be exempt in most countries. Methods applying the value added tax to financial services attempt to address this issue. This paper expands existing knowledge by integrating these methods into a single, unique formulation. A unified framework for generalising and characterising these approaches is presented, facilitating a new classification based on either shadow or implicit prices. Furthermore, two new variants of the mobile-ratio method for individual transactions, utilising either shadow prices or "explicit-made" implicit prices, are derived. These variants improve and address critiques from previous methods.

Keywords: value added tax, financial services, financial VAT, methods of taxation, optimality

1 INTRODUCTION

Value added tax (VAT) is among the most sophisticated and modern forms of taxation, offering significant benefits in terms of efficiency and revenue collection. Nonetheless, a VAT exemption for financial services remains in force in most countries due to the technical and practical difficulties of taxing them on a transaction-by-transaction basis. The European Commission (2021) has recently consulted financial entities regarding the potential elimination or reform of this exemption (Pardete, Santos and Leote, 2024). Although the exemption results in the under-taxation of the sector, unrecoverable input VAT also creates inefficiencies and economic losses for banks. Indeed, other countries, such as China and Brazil, have either recently implemented or are currently transitioning towards taxing financial services under a VAT regime.

Furthermore, this exemption leads to substantial losses in tax collection, undermines equity – given that the demand for many financial services is income elastic – and reduces efficiency. Over several decades, various authors (*inter alia*, Poddar and English, 1997; Huizinga, 2002; Zee, 2005) have sought to estimate the value added of financial services on a transaction-by-transaction basis (applied to financial transactions – flow variables – rather than financial accounts – stock variables) to facilitate the removal of this exemption. The most promising approach involves the efficient full taxation of both implicit and explicit fees and commissions. To the author's knowledge, this paper is the first to integrate the primary methods for applying VAT to financial services – specifically those that fully and accurately tax margin-based services – into a single, unified formula. A comprehensive framework for generalising and characterising these distinct methods is provided to reconcile differences in terminology and notation. Ultimately, this paper seeks to simplify the methods for applying VAT to financial services and serves as an instructive endeavour to enhance both the understanding and the subsequent improvement of these methods.

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

First, an overview of VAT is provided. VAT operates as a chain; a business applies the tax to its sales, and the customer pays this amount to the business, which is then required to remit it to the tax authorities. Conversely, the VAT paid on purchases (input VAT) and capital investments is paid by the business to its supplier and deducted from the output VAT. While this describes the process for an individual firm, the broader production cycle relies on the credit-invoice method.

However, this chain can be broken – most notably through exemptions, such as those applied to financial services in the majority of countries. The rationale for these exemptions is discussed later; first, the mechanics and immediate effects of exemptions are examined, drawing on the work of López-Laborda and Peña (2018) and the works they cite (*inter alia*, Grubert and Krever, 2012). When an exemption applies, input VAT cannot be credited; the business pays VAT to its supplier but cannot deduct against output VAT. This break in the chain triggers tax cascading, as unrecovered VAT inflates the final product price and incentivises vertical integration to avoid tax liabilities. Furthermore, this leads to the under-taxation of services for final consumers – distorting consumption patterns (Baydur and Yilmaz, 2021) – alongside with the over-taxation of registered business customers, who cannot recover the VAT embedded in their purchase prices (López-Laborda and Peña, 2018).

The primary reason for the exemption is the technical difficulty of applying VAT on a transaction-by-transaction basis. This stems largely from the dual nature of financial institutions as both suppliers (e.g., creditors) and buyers (e.g., depositors) of capital, even when the financial entity is the sole provider of the service. A significant challenge lies in isolating the implicit fee or margin – defined as the difference between interest receipts and payments – within a single transaction that may only involve interest receipts or payments.

Accounts-based approaches (utilising financial accounts rather than individual transactions) offer alternatives, such as the addition method used in Israel or Denmark, or the subtraction implemented in Italy via the regional production tax (*imposta regionale sulle attività produttive* – IRAP) unrelated to the national VAT, and proposed in Japan. However, these methods still interrupt the VAT chain unless universally applied. Consequently, the application of VAT to financial services remains an "unresolved issue" (Christiansen, 2017: 2), particularly regarding transaction-level implementation.

Another rationale for exempting financial services from VAT is the ongoing debate regarding whether such services yield direct consumer utility (Christiansen, 2017), whether they constitute final objects of utility (Boadway and Keen, 2003), or if only explicit fees should be subject to the tax (Jack, 2000). Grubert and Mackie (2000) go further, questioning whether an explicit tax on these services is justified at all. However, the point Grubert and Mackie originally make is that there is no consumption in borrowing and repaying with interest. At least insofar as the interest

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

is the risk-free rate of return, the borrowing and interest merely changes the time of consumption but is not consumption itself. In the paper written by Grubert and Krever (2012), they acknowledged the original view was incorrect as it failed to take into account that part of the interest paid was in addition to the risk-free rate of return interest, which merely changes the time of consumption, and its consideration for a supply that involves actual consumption of economic resources by the intermediary institution. This part of the interest, and only this part, he asserts should be subject to VAT.

The potential for positive tax revenue following the removal of the exemption is also questioned by examining the mechanisms of such taxation (Lockwood and Yerushalmi, 2019). Conversely, Auerbach and Gordon (2002) maintain that VAT should apply to inputs in the financial sector just as it does in other industries. Furthermore, taxing both explicit and implicit fees at a uniform rate avoids economic distortions (Peña, 2018). Weisbach (2019) recently reconciles the views of Grubert and Mackie (2000) with those of Auerbach and Gordon (2002), noting that both are predicated on the same logic: if financial services do not constitute consumption themselves, and if the resulting consumption can be taxed elsewhere, "there is no need to directly tax financial services" (p. 248).

Failing to levy VAT on both explicit and implicit financial fees also leads to distortions, as evidenced by recent empirical findings (Peña, 2026). Indeed, eliminating the exemption may yield beneficial effects on welfare – either by replacing distortive labour taxes (Buettner and Erbe, 2014) or when implemented at low general VAT rates (Baydur and Yilmaz, 2021). Such a move could also enhance tax collection and trade openness (Buettner and Erbe, 2014; Peña, 2018). Finally, removing the exemption may increase progressivity (Huizinga, 2002; Peña, 2018),[1] without the distortions to the financial sector (Peña, 2018) observed with financial transaction taxes.

This paper proposes two variants of the mobile-ratio (MR) approach (López-Laborda and Peña, 2018), which, when combined with other existing methods, offer potential solutions to the primary critiques of a regime of application of VAT to financial services. The MR method identifies the value added of each transaction by applying an accounts-based calculation to determine the financial margin. This margin is divided by the sum of interest receipts and payments to produce a ratio. This ratio is updated quarterly and applied to individual transactions to calculate the VAT liability.

[1] The view that the demand for many financial services is income-elastic and their most common users have high income, so the benefit of applying VAT to such services is progressively distributed, has been suggested by the previous authors, but if it refers to loan-deposit intermediary services by financial entities, it is undercut by the fact that higher income persons often borrow through registered enterprises for investments rather than consumption loans so they would be in a position to claim input tax credits and recover VAT completely. Also, to the extent lower-income individuals used borrowed funds for investment purposes, they would be overtaxed if they could not register and recover input tax.

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

The methods proposed herein also refine other theoretically robust systems, such as the Cash Flow (CF), the Modified Reverse-Charge (MRC) or the Hybrid System (HS). The CF method (Poddar and English, 1997) determines value added as the spread between transaction interest rates and an indexing rate derived from short-term bond yields. Conversely, the MRC (Zee, 2005) allows banks to choose the allocation between loans and deposits – a feature often criticised for its discretionary nature.

The General Mobile-Ratio methods introduced in this paper employ an implicit indexing rate. This is made explicit in the General Explicit variant, which mirrors the MR but utilises transaction-by-transaction calculations for the value added rather than accounts-based ones. Finally, while the HS method is similar to the MR, it applies specific ratios to distinct types of operations, whereas the MR uses a single ratio for all business activities within a period. In contrast, the General Shadow variant proposed here employs a unique value-added ratio for every transaction. Consequently, these proposed methods address existing deficiencies in the literature and resolve the specific criticisms associated with the MR.

The general formulation presented in this paper, supported by the underlying economic model and theoretical framework of optimisation programmes, enables the classification of the most promising current approaches. These are categorised by their level of application (transaction, type of transaction, or business) and the nature of taxation. The latter is approached in two ways: using implicit prices for pure interest via subtractions, or using shadow prices for interest via multiplicative ratios. Furthermore, this framework clarifies existing concepts, allowing for the proposal of two variants of a General Mobile-Ratio method at the transaction level.

Specifically, two alternatives are proposed: the General Explicit variant, involving the subtraction of explicit-made implicit pure interest rates; and the General Shadow variant, which applies the first ratio calculated on a transaction-by-transaction basis, an approach hitherto unaddressed in the literature. These methods address existing weaknesses and resolve criticisms of current "ideal" methods. For instance, unlike the CF or MRC methods, the proposed approach avoids issues regarding the choice of indexing rates; furthermore, unlike the HS or MR methods, it accounts for both prices and quantities. Numerical examples are provided to illustrate these concepts. The paper is organised as follows: section 2 reviews the literature providing the basis for VAT and VAT applied to financial services. Section 3 provides a common example of comparison. Sections 4 to 6 explain, respectively, the Cash-Flow (CF), Mobile-Ratio (MR) and Hybrid System (H) methods. Section 7 establishes the theoretical framework, presents the general formulation, integrating the most relevant taxation methods into a single expression, and proposes the two variants of the new General Mobile-Ratio method. Section 8 discusses the results, while section 9 concludes.

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

2 THE BASELINE FOR APPLYING VAT TO LOAN-DEPOSIT SERVICES

2.1 GENERAL VAT METHOD

Consumption is determined via the indirect subtraction method, in which sales are pY, p is the price and Y the output, the purchases are qx, with x being the input and q its price and I_G being the gross investment:

$$VAT = \tau pY - \tau qx - \tau I_G \tag{1}$$

where τ denotes the standard VAT rate. For the sake of simplicity, gross investments are hereafter excluded from this analysis, except where their inclusion is strictly necessary. The exemption status implies that while financial activities are not subject to output tax, financial entities are unable to recover input VAT incurred on non-financial purchases. Consequently, the revenue collected by the State comprises two elements: the VAT payable on financial services – which is nil due to the exemption – and the irrecoverable VAT paid on inputs. As the latter cannot be credited, it typically yields a positive tax collection. Therefore, the revenue accrued by the Public Administration through the VAT exemption of these services, calculated via the credit-invoice (or indirect subtraction) method, is represented by the following analytical expression:

$$VAT_{FS} = VAT_i + VAT_S = 0 + \tau qx = \tau qx \tag{2}$$

Here, VAT_{FS} denotes the Public Administration's VAT revenue on financial services, VAT_i represents the payable VAT by the financial institution, and VAT_S denotes the input VAT incurred on non-financial purchases. It is charged by the supplier, but not allowed to be creditable for the financial institution. The non-taxation of the financial margin within a VAT regime exacerbates the price increase in the first period relative to the second, a phenomenon primarily driven by the non-recoverability of input VAT.

2.2 COUNTRY EXPERIENCE

In Israel, VAT is applied to loan intermediary services through a general multi-stage system (Goldman, 2017: 1). Expenses related to goods or services at every stage of production, distribution, and marketing are taxed at a standard rate of 17%. While financial institutions that grant the loans are subject to a Wage and Profit Tax (also at 17%) on wages, salaries and earnings, their services remain VAT-exempt, meaning that entities cannot deduct input tax. Consequently, customers who take loans from financial institutions are unable to recover the tax paid by the financial institution, and deposits remain exempt.

In Argentina, the taxable base consists of the gross interest received by financial entities. While the standard VAT rate is 21%, certain loan intermediary services benefit from a reduced rate of 10.5%. Furthermore, Argentina imposes a provincial indirect tax on gross turnover, distinct from the national VAT (Teijeiro, 2017). Following a similar trajectory to Argentina (Peña, 2018), China replaced its Business

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

Tax regime with VAT on 1 May 2016, levying tax on the gross interest of loan supplies (Li and Krever, 2017). Although central bank and certain inter-bank loans are exempt, they are treated as exempt without credit, since input tax credits are denied. For most other loans, gross interest payments form the tax base; however, the inability of both borrowers and lenders to claim input tax credits effectively replicates the cascading effect of the former Business Tax.

Canada generally exempts financial services from the Goods and Services Tax/Harmonized Sales Tax (GST/HST). Entities providing both exempt and taxable supplies must apportion their input tax credit claims based on the extent to which inputs are utilised in taxable activities (Thang, 2017). Within the European Union, the granting of credit – where the full interest payment constitutes the taxable amount – is generally exempt without credit. Nevertheless, certain member states permit the "opt-to-tax" (O2T) for business-to-business (B2B) transactions (Englisch, 2017). Research suggests this method is more conducive to trade than alternative methods (Peña, 2018). In South Africa, a single-rate VAT applies to a broad base with minimal exemptions. While financial services are broadly exempt, explicit fees and commissions are excluded from this exemption and are thus subject to VAT. Input VAT may only be recovered if incurred for making taxable supplies, often determined by a turnover-based apportionment ratio (de Koker and Badenhorst, 2017).

Finally, New Zealand introduced a mechanism in 2005 allowing financial services suppliers to zero-rate supplies to specific GST-registered recipients (Pallot and Allen, 2017). New Zealand's system eliminates VAT (GST in their case) entirely on B2B transactions. The Australian system reduces the GST levied on financial intermediary supplies equally for B2C and B2B intermediary services but does not eliminate it for either. The Singapore system provides a complete return of input tax for acquisitions equal to the proportion of loans that are B2B loans. While the law does not specify that the banks must apply the remission of tax to their B2B loans, competitive pressures have led to this result, with the Singapore system consequently considered a de facto zero-rating for B2B services via a surrogate regime. Australia's GST model treats loan intermediary services as "input taxed" (exempt). To mitigate self-supply bias – which often disadvantages smaller competitors unable to bring services in-house – Australia allows a "reduced input tax credit" (typically 75%) for specific inputs used in loan intermediary activities (Krever and Teoh, 2017).

2.3 THE MOBILE-RATIO METHOD

López-Laborda and Peña (2018) propose a new method for levying VAT on financial services. This approach possesses desirable properties and is practically applicable. Termed the "mobile-ratio method", it utilises a quarterly updated proportion to allocate the financial institution's margin to individual transactions. Consequently, a near-complete taxation of the value added generated by financial services is achieved. This method is simple, neutral and easily administered. The authors' proposal is validated using data from the central banks of several advanced countries. Section 5 provides a further explanation of the method from a deeper perspective.

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

3 COMMON EXAMPLE OF COMPARISON

Following Peña (2018), a closed economy is assumed, featuring a banking sector (hereafter referred to as "banking") that performs identical financial operations of equal value. The economy further comprises households, the public sector, and firms. All financial operations provided by financial institutions are distributed equally between consumers and firms. We assume that across all methods, financial institutions and firms can pass through VAT applied to financial services to both consumers and other businesses. While the data represents a specific quarter, we assume that these transactions remain constant across all periods. Financial institutions are unable to recover input VAT on non-financial purchases due to the VAT exemption. These non-financial operations are detailed in table 1, based on an illustrative VAT rate of 10%. Non-financial purchases exclusive of VAT amount to €25, with €2.5 representing non-recoverable input VAT.

Table 1

Non-financial operations under the exemption

Non-financial operations	Gross amount	VAT rate	Input VAT
Non-financial purchases	25	0.1	2.5

Source: Author.

Financial operations comprise the gross transaction amount plus a fixed commission – which remains constant across all operations – and a variable commission. For the latter, it is assumed that non-deductible input VAT is distributed proportionally based on the sum of the absolute values of financial operations from previous years for each transaction, as illustrated in table 2.

Financial entities distributed total profits of €12 and incurred personnel costs of €125. Consequently, the sum of profits and salaries amounts to €137. Furthermore, the bank margin (difference between interest receipts and payments) is €150, while total commissions (fixed and variable) rise to €14.5. Non-financial purchases amount to €25, with €2.5 of non-deductible input VAT. In aggregate, the margin plus commissions, net of taxes (subtracting non-deductible VAT), also reaches a total of €137. It is worth highlighting that the variable commission is settled by the bank for passing through the non-deductible input VAT to the consumers (this is the reason for the coincidence of both in €2.5).

Table 2

Financial operations under the exemption

Financial operations	Gross amount	Variable commission	Fixed commission
Loan	4,000	1.21	3
Deposit	-4,000	1.21	3
Loan interest	200	0.06	3
Deposit interest	-50	0.02	3
Total	150	2.50	12

Source: Author.

The profits of banking are:

$$\Pi_B = I - G = D + IL - L + ID + VC + FC - VAT_{Input} - NFC - W = 4{,}000 + 200 - 4{,}000 - 50 + 2.5 + 12 - 2.5 - 25 - 125 = 12 \quad (3)$$

Where I represents income, G expenses, D deposit amount, IL loan interests, L loan amount, ID deposit interest, VC variable commissions, FC fixed commissions, VAT_{Input} input VAT, NFC non-financial purchases and W wages. The tax revenue from these operations, received by the Public sector, constitutes the unrecoverable input VAT:

$$Tax\ Collection\ (TC) = VAT_{Input} = 2.5 \quad (4)$$

The profits for customers (households and firms) are composed of:

$$\Pi_p = NCF - D - IL + L + ID - VC - FC + W = 25 - 4{,}000 - 200 + 4{,}000 + 50 - 2.5 - 12 + 125 = -14.5 \quad (5)$$

The aggregate profits from the whole economy reach (3) + (4) + (5):

$$\Pi_T = \Pi_B + TC + \Pi_p = 12 + 2.5 - 14.5 = 0 \quad (6)$$

Summarising, the derived economic situation appears in table 3.

Table 3

Economic situation by sectors and total economy with exemption

	Total revenue	Total expenses	Profits	Change
Financial entities	4,214.5	-4,202.5	12.0	0
Public sector	2.5	0	2.5	0
Private economy	4,200.0	-4,214.5	-14.5	0
Total			0	0
GDP (assets)			14.5	0

Source: Author.

The "change" column serves to illustrate the variation between each examined method and the exemption method (zero value in table 3 since it is the exemption case). The "Total" entry represents the aggregate of all assets and liabilities.

4 CASH-FLOW METHOD AND ANALOGOUS APPROACHES

4.1 FORMULATION, EXPLANATION AND CRITICISMS

Despite the global application of diverse approaches to taxing financial services, "there is no place where an ideal VAT model has been adopted in practice so far" (James, 2015, cited in Antonov, 2024: 138). Indeed, the VAT was rarely discussed more than seventy years ago, with the notable exceptions of France and select

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

academic or specialist texts (Keen and Lockwood, 2010). The transaction-by-transaction approach, frequently regarded in the literature as the most accurate, is the Cash-Flow (CF) method. Developed by Poddar and English (1997), this model differentiates between cash inflows and outflows, employing short-term bill rates as proxies for the "indexing rates" of implicit pure interest. Under this method, financial entities are considered to maximise the following financial margin:

$$Max\Pi_i = (\rho_{iL} - \varepsilon_{i,I})\, L + (\varepsilon_{i,I} - \rho_{iD})D \tag{7}$$

Here, i denotes the corresponding financial transaction, while L and D represent the loan and deposit amounts. Furthermore, ρ_{iL} and ρ_{iD} refer to their respective interest rates, classified as outflows and inflows depending on whether they constitute negative or positive cash flows for the business. The indexing interest rates, denoted by $\varepsilon_{i,I}$, serve as proxies for the risk-free rate. Under the Cash-Flow method, these rates are approximated by the short-term Treasury bill rate (Poddar and English, 1997) or the central bank deposit rate (Næss-Schmidt, Heebøll and Lund, 2016). Conversely, the modified reverse-charging proposal developed by Zee (2005) suggests that the choice of indicator should be left to the discretion of bankers, resulting in a discretionary indexing rate. The VAT collected generated by each financial transaction is formulated as follows[2]:

$$VAT_i = \tau \times [(\rho_{iL} - \varepsilon_{i,I})\, L + (\varepsilon_{i,I} - \rho_{iD})D - C] \tag{8}$$

with C representing the recoverable input and investment VAT.

As Ernst and Young (1998) illustrate, an advantage of the method is that it is applied to the actual interest rate for credits, allowing banks to refund bad credits. Although Mirrlees et al. (2011) provide a positive outlook on the method, they maintain that its neutrality depends on the chosen indexing rate – the "normal" rate of return – being accurate. Using an excessively high indexing rate would over-tax savings and under-tax loans, and vice versa, leading to distortions in financial activity patterns. There are several conceptual issues, such as the use of the same indexing rate for various operations regardless of whether there is a fixed or variable rate, maturity or other characteristics of the service (Ernst and Young, 1998). Furthermore, administrative issues exist; for instance, matching an indexing rate to the contract maturity of a financial service was considered "not feasible" by Ernst and Young (1998: 1). This method was analysed in practice by Ernst and Young in a 1998 pilot study for the European Commission, which tested the previous indexing rate and assessed the proposal of "a short-term inter-bank rate (either 1-month or 3-month rate)". While current Information and Communication Technologies (ICTs) could resolve the administrative issues (Slemrod and Velayudhan, 2022) that rendered the method unfeasible in the past – due to the then-limited state of enterprise systems and high computing costs – certain conceptual issues remain

[2] Notably, the current representations are abstractions of the underlying methods; the CF approach, for instance, accommodates the recognition of non-performing loans, unlike alternative methods.

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

unresolved and will be addressed in the following sections. These critiques suggest potential improvements to the approach, which will be analysed in section 7. Other methods in the literature include the Modified Reverse-Charging (MRC) approach developed by Zee (2005), which permits banks to choose the allocation of the estimated implicit fees among operations. A primary critique of this method is the discretionary nature of the bank's choice.

4.2 EXAMPLES

A tax revenue consistent with the theoretical formulation of the addition method is derived for aggregate effects. Assuming that the indexed interests – proxied by Treasury bill rates (Poddar and English, 1997), central bank deposit rates (Næss-Schmidt, Heebøll and Lund, 2016), and rates determined by banks (Zee, 2005) – amount to €100, the resulting financial operations align with the analogous results from table 2 – but applied to the cash-flow method, as illustrated in table 4.

TABLE 4

Application of the Peña (2018) example with the cash flow method

Financial operations	Gross amount	Fixed commission	Financial services VAT
Loan	4,000	3	0.3
Deposit	-4,000	3	0.3
Loan interest	200	3	(200-100+3)*0.1=10+0.3=10.3
Deposit interest	-50	3	(100-50+3)*0.1=5+0.3=5.3
Total	150	12	VAT to be paid: 15+0.3*4=16.2

Note: All variable commissions are zero.

Source: Author.

The CF method is applied to loans by subtracting the indexed interest (€100) from the loan interest (€200), with the resultant output then multiplied by the prevailing VAT rate. Furthermore, VAT applied to Explicit Fees and Commissions must be added to each transaction (e.g., €3×10% = €0.3). Regarding deposits, the deposit interest (€50) is subtracted from the indexing interest before the VAT rate is applied. Consequently, the economic position from table 4 remains equal to the one of the basic cash flow method. This is also consistent with the analogous tables from the next sections and methods, since there is identical calculated VAT received by the public sector.

A further different exercise for the implicit margin is provided in table 5, the format of which is maintained for the remaining methods. This example illustrates the mechanism employed in CF-type methods; it is adapted from the final table in Zee (2005), using the data given in the previous example from Peña (2018) at the 10% VAT rate, and also follows the procedure outlined in table 4 and formula (8).

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

Table 5
Example of Zee (2005) applied to the CF and MRC methods

Key	Assumptions	Method	Subject, VAT charged on:	Operation	Amount in €
a	Deposit and loan amounts = €4,000	CF	Loan interest	a*(c-e)*d	10
b	Deposit rate = 1.25%		Deposit interest	a*(e-b)*d	5
c	Loan rate = 5%	MRC	Loan interest	a*(c-(b+f))*d	7
d	VAT rate = 10%		Deposit interest	a*d*f	8
e	Indexing rate = 2.5%				
f	Allocation chosen by the bank = 2%				

Source: Author.

5 MOBILE RATIO

5.1 FORMULATION, EXPLANATION AND CRITICISMS

The Mobile-Ratio (MR) method, proposed by López-Laborda and Peña (2018), applies a similar ratio to the one of the Hybrid System Method described in section 6; however, it is updated quarterly and utilises a single ratio per business for each period. Consequently, the MR method establishes a distinct proportion for every firm.

Theorem 1. Maximisation of profits in a business applying the MR method

$$\underset{\{L,D\}}{Max}\ \Pi = (\rho_L - \varepsilon_k)L + (\varepsilon_k - \rho_D)D \quad s.t.: \rho_L L + \rho_D D = TVI \Rightarrow \begin{cases} \rho_k := -\lambda = \dfrac{\rho_L - \rho_D}{\rho_L + \rho_D} \\ \varepsilon_k = \dfrac{2\rho_L \rho_D}{\rho_L + \rho_D} \end{cases} \tag{9}$$

Proof: See appendix.

The preceding maximisation expression is similar to that of the previous methods (7); however, it incorporates the Total Value of Interests (TVI). Defined by López-Laborda and Peña (2018) as the sum of interest receipts and interest payments, the MR approach considers the business in its entirety rather than assessing each transaction in isolation. The variable ρ represents the "mobile ratio" according to the authors' terminology. Within this model, it may be interpreted as a "shadow price" (a Lagrange multiplier, with a negative sign), signifying the approximate increase in the financial margin when one unit of TVI is raised. This method attains the primary desirable features of the transaction-based methods while maintaining the advantages of accounts-based methods. Furthermore, explicit fees are taxed via the credit-invoice method, ensuring the VAT chain remains intact as input VAT is recoverable.

The VAT collected for each business is determined as follows:

$$VAT = \tau \times [\rho_k \rho_L L + \rho_k \rho_D D - C] \tag{10}$$

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

One critique, related to insurance services (Altenburger, 2022), suggests that because the original formulation of the mobile-ratio method does not account for capital amounts, insurance services – such as non-life insurance – are not adequately addressed. Additionally, insurance providers often experience periods with negligible expenditure, particularly during their initial phase of operations. This can lead to temporal mismatches between income and expenditures, which the mobile-ratio method may incorrectly reflect.[3]

A second critique posits that loan interests are over-weighted in the obtainment of the payable tax amount. In this regard, the Tax Working Group Secretariat (WGS, 2018, fn. 10) of New Zealand, while noting the method's potential simplicity, states that it "appears to attribute greater amounts of the financial service to borrowers rather than depositors". This suggests a higher levy on loans than is perhaps appropriate, stemming from the "approximate" calculation of the value added on a transaction-by-transaction basis.

A further critique is offered by López-Laborda and Peña (2018), who confirm a trade-off between accuracy and neutrality – as well as between accuracy and feasibility. They argue that the cash flow method is conceptually accurate but "not feasible" in practice (Ernst and Young, 1998), at least prior to the advent of modern ICTs. The primary barrier to feasibility is the requirement to calculate the "pure" interest rate – the indexing rate – for every transaction to subtract it from the corresponding interest. While the MR method offers greater feasibility, it suffers from a degree of inaccuracy since the ratio used to estimate the value added remains constant across all transactions within a business.

5.2 EXAMPLES

A tax revenue is obtained that is equivalent to using the theoretical formulation of the cash flow method in aggregate. However, results for loans and deposits vary individually; this highlights that, within the CF framework, the selection of the indexation interest rate requires a degree of discretion due to computational difficulties. In contrast, the underlying indexation interest rate derived from the proposed method is highly representative of the short-term bond yields, consistent with the findings of López-Laborda and Peña (2018).

According to their formulation, the implicit pure interest would be €80, calculated as $2\times200\times50/(200+50)=80$. Given that the mobile ratio is 60%, derived from $(200-50)/(200+50)$, the financial operations align with the data presented in table 2 (see table 6).

[3] I would want to thank Ben Lockwood for suggesting this idea to me at the OFS Workshop of Indirect Taxes, held on 30-31 March 2017, where the paper about the mobile-ratio method was presented.

Table 6
Example of Peña (2018) with the MR method

Financial operations	Gross amount	Fixed commission	Financial services VAT
Loan	4,000	3	0.3
Deposit	-4,000	3	0.3
Loan interest	200	3	(200*0.6+3)*0.1=(200-80+3)*0.1=12+0.3=12.3
Deposit interest	-50	3	(50*0.6+3)*0.1=(80-50+3)*0.1=3+0.3=3.3
Total	150	12	VAT to be paid: 15+0.3*4=16.2

Note: All variable commissions are zero.

Source: Author.

This mobile-ratio method applies the ratio – calculated in aggregate via the addition method using data from the previous quarter – to the interest on loans and deposits. This approach is equivalent to the CF, utilising a financial services gravity equation for the indexation factor. VAT is levied on this sum, in addition to commissions. Example 5, sourced from the final table in Zee (2005), is illustrated in table 7 as applied to the MR method, following the procedure outlined in table 5, from which the hypotheses and given data are the same in this common example throughout the paper in order to make the methods comparable.

Table 7
Example of Zee (2005) applied to the MR method

Key	Assumptions	Method	Subject, VAT charged on:	Operation	Amount in €
a	Deposit and loan amounts = €4,000	MR	Loan interest	a*c*d*g	12
b	Deposit rate = 1.25%		Deposit interest	a*b*d*g	3
c	Loan rate = 5%				
d	VAT rate = 10%				
e	Indexing rate = 2.5%				
f	Allocation chosen by the bank = 2%				
g	Business mobile ratio = 0.6				

Source: Author.

6 HYBRID SYSTEM

6.1 FORMULATION, EXPLANATION AND CRITICISMS

Another proposal is the Hybrid System (HS), put forward by Bird and Gendron (2005), which consists of applying a ratio of the value added per type of transaction. A previous version, initially developed by Bakker and Chronican (1985) and termed "separate tax rate method", was originally intended to apply only to margin-based operations. However, Bird and Gendron (2005) extended this method to explicit fees and commissions, developing the hybrid system explained here.

Theorem 2. Profit maximisation in a business applying the Hybrid System

$$\left.\begin{array}{c} Max\ \Pi_L = \left(\rho_L - \varepsilon_j\right) L \\ s.t.: \rho_L L = IR \end{array}\right\} \Rightarrow \rho_{j,L} := -\lambda_L = \frac{\rho_L - \varepsilon_j}{\rho_L}$$
$$\left.\begin{array}{c} Max\ \Pi_D = \left(\varepsilon_j - \rho_D\right) D \\ s.t.: \rho_D D = IP \end{array}\right\} \Rightarrow \rho_{j,D} := -\lambda_D = \frac{\varepsilon_j - \rho_D}{\rho_D} \qquad (11)$$

Proof: See appendix.

As can be observed, in contrast to the two previous methods, a single proportion is applied to each type of financial transaction; this indicates a transaction-type level analysis by *j*. Therefore, the proportion differs from loans and deposits, resulting in two maximisation programmes within the same business by distinguishing between market sectors. Specifically, there is a maximisation of profits for each type of financial transaction. The proportion also assumes the role of a shadow price (minus the Lagrange multipliers); in this case, it represents the increase in, respectively, interest receipts *IR* and interest payments *IP*.

The VAT collected for each business is as follows:

$$VAT = \tau \times [\rho_{j,L}\rho_L L + \rho_{j,D}\rho_D D - C] \qquad (12)$$

The method has several disadvantages. First, it requires some approximations. Second, it incurs higher administrative and compliance costs than aggregate methods, such as addition or subtraction. Third, as Gendron (2008) points out, obtaining data on the value added of each type of transaction – as well as the share of value added for each transaction relative to its total volume – can be complex and costly. Furthermore, it is occasionally difficult to distinguish between different types of transactions.

6.2 EXAMPLES

A tax collection is obtained that is equivalent in aggregate to that obtained via the theoretical formulation of the cash flow method and of the mobile ratio. However, the results for loans and deposits change individually due to the specific parameters of the example.[4] Simplifying the formulation of Bird and Gendron (2005), the proportion of interest is calculated using the same pure interest of €80 as in the previous method. Thus, the proportion for loan interest would be 60%, calculated as (200-80)/200, which is identical to that for deposit interests, (80-50)/50. These operations are shown in table 8 below, following the operations from table 2.

[4] Nonetheless, missadjustments can appear in real life, since they are approximations.

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

TABLE 8

Application of the Peña (2018) example under the hybrid system

Financial operations	Gross amount	Fixed commission	Financial services VAT
Loan	4,000	3	0.3
Deposit	-4,000	3	0.3
Loan interest	200	3	(200*(200-80)/200+3)*0.1=12+0.3=12.3
Deposit interest	-50	3	(50*(80-50)/50+3)*0.1=3+0.3=3.3
Total	150	12	VAT to be paid: 15+1.2=16.2

Note: All variable commissions are zero.

Source: Author.

This hybrid system method applies a ratio per transaction. Knowing that it is difficult to distinguish among different types of transactions, in this paper it is proposed to apply the ratio given by equation (11). Following the methodology established by Zee (2005), an illustrative example is provided in table 7, adapted for the Hybrid System in accordance with the procedures detailed in table 9.

TABLE 9

Zee's (2005) example applied to the Hybrid System

Key	Assumptions	Method	Subject, VAT charged on:	Operation	Amount in €
a	Deposit and loan amounts = €4,000	HS	Loan interest	a*c*d*h	12
b	Deposit rate = 1.25%		Deposit interest	a*b*d*i	3
c	Loan rate = 5%				
d	VAT rate = 10%				
e	Indexing rate = 2.5%				
f	Allocation chosen by the bank = 2%				
g	Business mobile ratio = 0.6				
h	Loan-type ratio = 0.6				
i	Deposit-type ratio = 0.6				

Source: Author.

7 PROPOSAL: GENERAL FORMULA AND GENERAL MOBILE-RATIO METHOD (G) VARIANTS

7.1 FORMULATION

Building on the development of the previous theoretical framework, this section proposes the General Mobile-Ratio Method (G). This framework calculates the value added on a transaction-by-transaction basis rather than applying the MR method as in the basic, original method. The new proposal integrates the use of mobile ratios with explicit indexed (implicit pure interest) rates. The classification

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

of these methods is based on a new criterion: whether the system accounts for value added by subtracting implicit pure – or indexed rates – (Implicit Price Systems, IPS); or by applying shadow prices as proportions of interest rates (Shadow Price Systems, SPS). The proposed formula for VAT collection is as follows:

$$VAT_k = \tau \times [(\rho_{\mu,SPS}\rho_{\mu,L} - \varepsilon_{\mu,IPS})L + (\varepsilon_{\mu,IPS} - \gamma\rho_{\mu,SPS}\rho_{\mu,D})D - C] \tag{13}$$

The specific cases are compared in the subsequent section. The subscript μ denotes the level of VAT collection, with $\mu \in \mathbb{N}^+$ and $\mu = \{i, j, k\}$: transaction-by-transaction (denoted by i, as in CF and proxies), business-level (k, as in MR), or transaction-type level (by j, as in HS). Furthermore, IPS methods assume $\rho_{k,SPS} = \gamma = 1$ (encompassing CF and similar approaches), whereas SPS methods utilise $\varepsilon_{k,IPS} = 0$ and $\gamma = -1$ (including MR and HS), as they are additive rather than subtractive. The General Explicit (GE) and General Shadow (GS) variants are derived by applying IPS and SPS principles on a transaction-by-transaction basis. These utilise transaction-specific pure rates and mobile ratios derived from the optimisation problem presented in Theorem 3.

Theorem 3. Profit maximisation in a business applying the G method

$$\begin{aligned} Max\ \Pi_i &= (\rho_{iL} - \varepsilon_i)L + (\varepsilon_i - \rho_{iD})D \\ s.t.: \rho_{iL}L + \rho_{iD}D &= TVI_i \end{aligned} \Rightarrow \begin{cases} \rho_i := -\lambda_i = \dfrac{\rho_{iL} - \rho_{iD}}{\rho_{iL} + \rho_{iD}} \\ \varepsilon_i = \dfrac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} \end{cases} \tag{14}$$

Proof: See appendix.

The individual mobile ratio functions as a shadow price (a negative Lagrange multiplier) with an interpretation similar to that in the MR method, albeit applied to an individual transaction. It represents the increase in the financial margin of a transaction resulting from a 1% increase in the total interest value of that transaction. The explicit pure interest rate is expressed as a rate rather than an absolute amount. These two variants offer several advantages: first, accuracy, by applying calculations to capital amounts at the transaction level, they improve upon the accuracy of the MR method. Second, insurance treatment, since they allow for the precise treatment of non-life insurance, which is claimed to be problematic under the MR method (Altenburger, 2022). Third, consistency, by eliminating the need to distinguish between transaction types (as in HS) and avoiding the allocation issues inherent in CF – such as when the indexing rate falls outside the spread between lending and deposit rates. Finally, objectivity, since they remove the discretionary estimation of rates permitted under the MRC approach.

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

The GE (15) and GS (16) variants are expressed as:

$$VAT_i = \tau \times \left[\left(\rho_{iL} - \varepsilon_i \right) L + \left(\varepsilon_i - \rho_{iD} \right) D - C \right] = \tau \times \left[\left(\rho_{iL} - \frac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} \right) L + \left(\frac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} - \rho_{iD} \right) D - C \right] \quad (15)$$

$$VAT_i = \tau \times \left[\rho_i \rho_{iL} L + \rho_i \rho_{iD} D - C \right] = \tau \times \left[\frac{\rho_{iL} L - \rho_{iD} D}{\rho_{iL} L + \rho_{iD} D} \rho_{iL} L + \frac{\rho_{iL} L - \rho_{iD} D}{\rho_{iL} L + \rho_{iD} D} \rho_{iD} D - C \right] \quad (16)$$

where ε_i and ρ_i represent the transaction-specific explicit pure rate and mobile ratio, respectively.

7.2 EXAMPLES

While the aggregate tax collection remains equivalent to previous methods, these variants provide greater granularity, achieving near-exact accuracy per transaction rather than per business (López-Laborda and Peña, 2018). Under the standard MR method, the implicit pure interest would be €80, 2×200×50/(200+50), and the individual mobile ratio would be 60%, (200-50)/(200+50). However, for both GE and GS, these values are calculated independently for each transaction, rather than being derived from the aggregate accounts of the business. The financial operations are detailed in table 2; specific applications are shown in table 10 (GE) and table 11 (GS).

Table 10

Application of the Peña (2018) example under the GE method

Financial operations	Gross amount	Fixed commission	Financial services VAT
Loan	4,000	3	0.3
Deposit	-4,000	3	0.3
Loan interest	200	3	(200-2*200*50/(200+50)+3)*0.1=12+0.3=12.3
Deposit interest	-50	3	(2*200*50/(200+50)-50+3)*0.1=3+0.3=3.3
Total	150	12	VAT to be paid: 15+0.3*4=16.2

Note: All variable commissions are zero.

Source: Author.

Table 11

Application of the Peña (2018) example under the GS method

Financial operations	Gross amount	Fixed commission	Financial services VAT
Loan	4,000	3	0.3
Deposit	-4,000	3	0.3
Loan interest	200	3	(200*(200-50)/(200+50)+3)*0.1=12+0.3=12.3
Deposit interest	-50	3	(50*(200-50)/(200+50)+3)*0.1=3+0.3=3.3
Total	150	12	VAT to be paid: 15+0.3*4=16.2

Note: All variable commissions are zero.

Source: Author.

The following example, adapted from Zee (2005) and presented in table 12, applies the GE and GS methods according to the procedure outlined in tables 10 and 11.

Table 12

Application of the Zee (2005) example to GE and GS methods

Key	Assumptions	Method	Subject, VAT charged on:	Operation	Amount in €
a	Deposit and loan amounts = €4,000	GE	Loan interest	a*(c-2*b*c/(b+c))*d	12
b	Deposit rate = 1.25%		Deposit interest	a*(2*b*c/(b+c)-b)*d	3
c	Loan rate = 5%	GS	Loan interest	a*c*(c-b)/(b+c)*d	12
d	VAT rate = 10%		Deposit interest	a*b*(c-b)/(b+c)*d	3
e	Indexing rate = 2.5%				
f	Allocation chosen by the bank = 2%				
g	Business mobile ratio = 0.6				
h	Loan-type ratio = 0.6				
i	Deposit-type ratio = 0.6				
2*b*c/(b+c)	Explicit-made pure rate = 2%				
(c-b)/(b+c)	Shadow ratio = 0.6				

Source: Author.

Table 13 provides a numerical example to show the policy implementation of the GE and GS methods. The table details the mechanisms for applying the chosen policy instrument – the tax rate – under the mobile-ratio method. Here, the tax rate is positive for interest receipts and negative for interest payments; this reflects that bank receipts and payments represent costs and revenues for the consumer, respectively. As shown in the preceding table, the tax paid to the public administration is calculated by multiplying the interest by the explicit pure interest rate and the individual-transaction mobile ratio (GS) from equation (16), and subsequently applying the VAT rate (10% in the example). Total interest receipts and payments are assumed to remain constant before and after taxes.

Since VAT is a consumption tax levied on households, the rationale is that it must result in a reduction of household income. It is important to note that interest receipts by and interest payments to a bank constitute, respectively, negative and positive income for households. Consequently, in the former case (a payment), the tax is added, whereas in the latter, it is subtracted by the bank. In this arrangement, the bank acts as the ultimate payee and administrator, transferring the tax from households to the public administration.

Table 13

Illustration of tax application for GE and GS methods

Agent	Bank			$\rho_D D$	Household	
Case	**Transaction**	**Sign**	**Amount**	**Direction**	**Transaction**	**Amount**
Without tax or zero tax rate	Gross interest payments	–	0.2	→	0.2	Receives
With tax	Tax payments to the public administration	–	**GE**: ((2*0.3*0.2)/(0.3+0.2)-0.2)*0.1)=0.004 **GS**: 0.2*0.1*(0.3-0.2)/(0.3+0.2)=0.004			
	Net interest payments to the household	–	0.2-0.004=0.196	→	0.196	Receives
Agent	**Bank**			$\rho_L L$	**Household**	
Case	**Operation**	**Sign**	**Amount**	**Direction**	**Operation**	**Amount**
Without tax or zero tax rate	Gross interest receipts	+	0.3	←	0.3	Pays
With tax	Tax payments to the public administration	–	**GE**: (0.3-(2*0.3*0.2)/(0.3+0.2))*0.1=0.006 **GS**: 0.3*0.1*(0.3-0.2)/(0.3+0.2)=0.006			
	Net interest receipts from the household	+	0.3-0.006=0.294	←	0.294	Pays

Source: Author.

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

8 DISCUSSION

8.1 PARTICULAR CASES AND COMPARISONS DERIVED FROM THE GENERAL FORMULA

Table 14 summarises the primary theoretical differences encountered when applying the various existing methods (8-12) and the proposed methods (15, 16), in light of the general formula (13) introduced in this study.

TABLE 14

The general formula: comparison between margin-based methods

General formula: $VAT_k = \tau \times \left[\left(\rho_{\mu,SPS} \rho_{\mu L} - \varepsilon_{\mu,IPS} \right) L + \left(\varepsilon_{\mu,IPS} - \gamma \rho_{\mu,SPS} \rho_{\mu D} \right) D - C \right]$ (13)

Method	IPS/ SPS	Level (μ)	Value of the parameters	Formula	Equation
CF	IPS	i: transactions	$\rho_{\mu,SPS} = \gamma = 1$	$VAT_i = \tau \times [(\rho_{iL} - \varepsilon_{i,I}) L + (\varepsilon_{i,I} - \rho_{iD}) D - C]$	(8)
MR	SPS	k: business	$\varepsilon_{k,IPS} = 0$ & $\gamma = -1$	$VAT = \tau \times [\rho_k \rho_L L + \rho_k \rho_D D - C]$	(10)
HS	SPS	j: transaction-type	$\varepsilon_{k,IPS} = 0$ & $\gamma = -1$	$VAT = \tau \times [\rho_{j,L} \rho_L L + \rho_{j,D} \rho_D D - C]$	(12)
GE	IPS	i: transactions	$\rho_{\mu,SPS} = \gamma = 1$	$VAT_i = \tau \times \left[(\rho_{iL} - \varepsilon_i) L + (\varepsilon_i - \rho_{iD}) D - C \right] = \tau \times \left[\left(\rho_{iL} - \frac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} \right) L + \left(\frac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} - \rho_{iD} \right) D - C \right]$	(15)
GS	SPS	i: transactions	$\varepsilon_{k,IPS} = 0$ & $\gamma = -1$	$VAT_i = \tau \times \left[\rho_i \rho_{iL} L + \rho_i \rho_{iD} D - C \right] = \tau \times \left[\frac{\rho_{iL} L - \rho_{iD} D}{\rho_{iL} L + \rho_{iD} D} \rho_{iL} L + \frac{\rho_{iL} L - \rho_{iD} D}{\rho_{iL} L + \rho_{iD} D} \rho_{iD} D - C \right]$	(16)

Source: Author.

Table 14 provides a reminder of the general formula proposed for all main margin-based methods (13). The first column indicates the method applied, while the second reflects the new classification of these methods regarding the use of Implicit Price Systems (IPS) or Shadow Price Systems (SPS). Specifically, IPS covers all CF methods, their variants, and the previously proposed GE method. The remaining approaches (MR, HS and GS) are classified as SPS.

The third column accounts for the level of tax application, denoted by μ in the general formula (13). This takes the value *i* when applied differently for each financial transaction (CF, GE, and GS methods), the parameter *k* when applied per business (MR), and the value *j* when there is a distinct proportion for each type of transaction. The fourth column denotes the parameters for each method. These values remain consistent for IPS (using a subtractive approach, this is the rationale for $\gamma = 1$, hence maintaining the minus sign in formula 13), while differing on $\varepsilon_{k,IPS}$ for SPS (using a multiplicative-additive way, this is the reason for $\gamma = -1$, thus adding with positive sign in formula 13). The

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

general formula achieves different specifications for $\rho_{k,IPS}$. Finally, the fifth and sixth columns display the respective formulae and equations for the margin-based methods analysed, representing specific cases of the general formula (13).

8.2 EXAMPLES AND COMPARISON

Following the example in section 3, table 15 illustrates the economic situation resulting from the margin-based methods analysed above. This is compared with the VAT exemption for financial services shown in table 3 (section 3); this comparison is detailed in the final column, "Change" (with respect to table 3).

Banking profits remain as defined in expression (1), at €12. The public revenue generated by the Administration from these operations is:

$$\begin{aligned} TC &= 21\% * (\mathrm{DL} + \mathrm{IL} - \mathrm{L} - \mathrm{ID}) + \mathrm{VAT}_{\mathrm{Input}} - \mathrm{VAT}_{\mathrm{Input}} = \\ &10\% * (4{,}000 + 200 - 4{,}000 - 50) + 2.5 - 2.5 = 10\% * 150 = 15 \end{aligned} \tag{17}$$

profits for customers (consumers and companies) are:

$$\begin{aligned} \Pi_P &= \mathrm{CNF} - \mathrm{D} - \mathrm{IP} + \mathrm{P} + \mathrm{ID} - \mathrm{CV} - \mathrm{CF} + \mathrm{W} + \mathrm{TC} + \mathrm{VAT}_{\mathrm{Inputt}} = \\ &25 - 4{,}000 - 200 + 4{,}000 + 50 - 2.5 \ - 12 + 125 - 15 + 2.5 \\ &- 14.5 - 15 + 2.5 = -27 \end{aligned} \tag{18}$$

Total economic profits are the sum of (1), (2), and (3):

$$\Pi_T = \Pi_B + TC + \Pi_P = 12 + 15 - 27 = 0 \tag{19}$$

The resulting economic landscape is summarised in table 15.

Tables 16 and 17 show the intermediary operations for determining the final VAT payable under the analysed margin-based (MB) methods; the former addresses existing literature, while the latter focuses on the proposed methods. The tables illustrate a diverse range of existing and proposed methods, all yielding the same aggregate result. They provide a practical and illustrative complement to the general formula, reflecting current practices while facilitating technical comparisons and the derivation of the aforementioned proposals. Finally, table 18 summarises the procedures of all the analysed MB methods using the example proposed by Zee (2005), further streamlining technical comparisons.

TABLE 15

Peña's (2018) example, calculation of the VAT payable and economic situation

	Total revenue	Total expenses	Profits	Change
Financial entities	4,214.5	-4,202.5	12	0
Public sector	17.5	-2.5	15	12.5
Private economy	4,200.0	-4,227.0	-27	-12.5
Total			0	0
GDP (assets)			27	12

Source: Author.

Table 16

Peña's (2018) example, financial operations: CF method and existing MB methods

Financial operations	Gross amount	Fixed commission	CF: financial services VAT	MR: financial services VAT	HS: financial services VAT
Loan	4,000	3	0.3	0.3	0.3
Deposit	-4,000	3	0.3	0.3	0.3
Loan interest	200	3	(200-100+3)*0.21=21.63	(200*0.6+3)*0.21=(200-80+3)*0.21=25.83	(200*(200-80)/200+3)*0.1=12.3
Deposit interest	-50	3	(100-50+3)*0.21=11.13	(50*0.6+3)*0.21=(80-50+3)*0.21=6.93	(50*(80-50)/50+3)*0.1=3.3
Total	150	12	VAT to be paid: 34.02	VAT to be paid: 34.02	VAT to be paid: 16.2

Source: Author.

Table 17

Peña's (2018) example, financial operations: proposed MB methods

Financial operations	Gross amount	Fixed commission	GE: financial services VAT	GS: financial services VAT
Loan	4,000	3	0.3	0.3
Deposit	-4,000	3	0.3	0.3
Loan interest	200	3	(200-2*200*50/(200+50)+3)*0.21=25.83	(200*(200-50)/(200+50)+3)*0.1=12.3
Deposit interest	-50	3	(2*200*50/(200+50)-50+3)*0.21=6.93	(50*(200-50)/(200+50)+3)*0.1=3.3
Total	150	12	VAT to be paid: 34.02	VAT to be paid: 16.2

Source: Author.

Table 18

Example of Zee (2005) applied to the GE and GS methods

Method	Subject, VAT charged on:	Operation	Amount in €
CF	Loan interest	a*(c-e)*d	10
	Deposit interest	a*(e-b)*d	5
MRC	Loan interest	a*(c-(b+f))*d	7
	Deposit interest	a*d*f	8
MR	Loan interest	a*c*d*g	12
	Deposit interest	a*b*d*g	3
HS	Loan interest	a*c*d*h	12
	Deposit interest	a*b*d*i	3
GE	Loan interest	a*(c-2*b*c/(b+c))*d	**12**
	Deposit interest	a*(2*b*c/(b+c)-b)*d	**3**
GS	Loan interest	a*c*(c-b)/(b+c)*d	**12**
	Deposit interest	a*b*(c-b)/(b+c)*d	**3**

Source: Author.

9 CONCLUDING REMARKS

This paper addresses various issues regarding the methods of taxing financial services under the VAT. First, I develop a theoretical framework from where the primary methods in the literature are derived, enabling the formulation of new approaches. Additionally, I provide a correspondence identified between the "mobile ratio" proposed by López-Laborda and Peña (2018) and the "separate tax rates" suggested by Bird and Gendron (2005). Second, I propose a unified mathematical expression that is employed to synthesise the most efficient methods found in the literature. Finally, I address some current critiques of existing ideal approaches and propose a new method with two variants. These improve upon existing models by applying individual ratios to each transaction (the General Shadow variant) or, equivalently, an explicit, pure interest rate – excluding fees and risk – per transaction (the General Explicit variant). A further refinement considers interest rates and transaction volumes to address concerns regarding start-up businesses and non-life insurance. Moreover, I address potential misallocations of value added between loans and deposits by developing a ratio for each transaction type, rather than a single ratio per business (as in the Mobile Ratio method, MR) or per transaction at a national level (as in the hybrid method). Finally, the "G method" is shown to achieve greater accuracy than the MR and higher feasibility than the Cash Flow (CF).

The proposed methods are better alternatives to apply VAT to financial services than current practices and methods for the following reasons. First, the proposals are generally more accurate with regard to the implicit pure interest, because they use (implicitly or explicitly) a theoretically derived formula to allocate the financial margin among borrowers and depositors rather the inaccurate approximated indexing rates used in the Cash-Flow (CF) and analogous methods. They are also more accurate than Hybrid System (HS) or Mobile-Ratio (MR) methods since the value added is calculated in a more granular way, transaction-by-transaction instead of

per type of transaction or business, respectively. Second, the proposals are more feasible than others since they do not require the difficult process of distinguishing among different types of transactions (HS) nor approximated indexing rates that lead to technical problems and misallocations (CF).

Replacing the EU-harmonised VAT exemption G method for financial services could serve as alternative to be applied in the recent European tax reform that is planned on the application of VAT to financial services. This could be achieved by broadening the VAT base and applying the proposed General variants of the Mobile-Ratio method. Given that Gale (2020) recently proposed the MR and CF methods for the US, these variants also offer a solution for developing countries – such as those in Africa (Cnossen, 2019) – that may lack the Information and Communication Technologies (ICT) infrastructure required for complex methods (Buljan, 2021), such as the CF.

Disclosure statement

The authors have no conflicts of interest to declare.

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

No generative AI or AI-assisted tools were used in the preparation of this manuscript.

REFERENCES

1. Altenburger, O. A., 2022. Applying VAT to financial services: Is the new "mobile-ratio method" adequate for insurance? *Zeitschrift für die gesamte Versicherungswissenschaft*, 111(3), pp. 353-360. https://doi.org/10.1007/s12297-022-00532-5
2. Antonov, L., 2024. Value Added Tax on Financial Services in the EU: The Complete Story. *Review of European and Comparative Law*, 58(3), pp. 53-71. https://doi.org/10.31743/recl.17166
3. Auerbach, A. J. and Gordon, R. H., 2002. Taxation of Financial Services under a VAT. *American Economic Review*, 92(2), pp. 411-416. https://doi.org/10.1257/000282802320191714
4. Bakker, C. and Chronican, P., 1985. *Financial Services and the GST: a discussion paper*. Wellington, NZ: Victoria University Press for the Institute for Policy Studies.
5. Baydur, I. and Yilmaz, F., 2021. VAT Treatment of the Financial Services: Implications for the Real Economy. *Journal of Money, Credit and Banking*, 53(8), pp. 2167-2200. https://doi.org/10.1111/jmcb.12780
6. Bird, R. and Gendron, P. P., 2005. *VAT Revisited, A New Look at the Value Added Tax in Developing and Transitional Countries*. Toronto: International Tax Program, University of Toronto.
7. Boadway, R. and Keen, M., 2003. Theoretical perspectives on the taxation of capital income and financial services: A survey. In: P. Honohan, ed. *Taxation of Financial Intermediation: Theory and Practice for Emerging Economies*. New York: World Bank and Oxford University Press, pp. 31-80.
8. Buettner, T. and Erbe, K., 2014. Revenue and welfare effects of financial sector VAT exemption. *International Tax and Public Finance*, 21, pp. 1028-1050. https://doi.org/10.1007/s10797-013-9297-5
9. Buljan, A., 2021. Modernizing VATs in Africa. *Public Sector Economics*, 45(1), pp. 171-176. https://doi.org/10.3326/pse.45.1.6
10. Christiansen, V., 2017. Indirect Taxation of Financial Services. *CESifo Working Paper Series,* No. 6675. https://doi.org/10.2139/ssrn.3058577
11. Cnossen, S., 2019. *Modernizing VATs in Africa*. Oxford University Press.
12. de Koker, A. and Badenhorst, G., 2017. Loan Intermediary Services: South Africa. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 169-186.
13. Englisch, J., 2017. Loan Intermediary Services: European Union. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 123-138.
14. Ernst and Young, 1998. *The TCA System – A Detailed Description. Taxation and Customs Union, Reports and Studies Commissioned for the European Commission*. Brussels: Ernst and Young.
15. European Commission, 2021. *Public consultation "VAT rules for financial and insurance services today and tomorrow" Factual summary report*.

16. Gale, W. G., 2020. Raising revenue with a progressive value-added tax. In: E. Moss, R. Nunn and J. Shambaugh, eds. *Tackling the Tax Code: Efficient and Equitable Ways to Raise Revenue*. The Hamilton Project, pp. 43-48.
17. Gendron, P. P., 2008. VAT Treatment of Financial Services: Assessment and Policy Proposal for Developing Countries. *Bulletin for International Taxation*, 62(11), pp. 494-507.
18. Goldman, D., 2017. Loan Intermediary Services: Israel. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 139-158.
19. Grubert, H. and Krever, R., 2012. VAT and Financial Services: Competing Perspectives on What Should Be Taxed. *Tax Law Review*, 65, pp. 199-239.
20. Grubert, H. and Mackie, J., 2000. Must financial services be taxed under a consumption tax? *National Tax Journal*, 53(1), pp. 23-40. https://doi.org/10.17310/ntj.2000.1.02
21. Huizinga, H., 2002. A European VAT on financial services? *Economic Policy*, 17(35), pp. 497-534. https://doi.org/10.1111/1468-0327.00095
22. Jack, W., 2000. The treatment of financial services under a broad-based consumption tax. *National Tax Journal*, 53(4), pp. 841-851. https://doi.org/10.17310/ntj.2000.4.03
23. James, K., 2015. *The Rise of the Value-Added Tax*. Cambridge: Cambridge Tax Law Series.
24. Keen, M. and Lockwood, B., 2010. The value added tax: Its causes and consequences. *Journal of Development Economics*, 92(2), pp. 138-151. https://doi.org/10.1016/j.jdeveco.2009.01.012
25. Krever, R. and Teoh, J., 2017. Loan intermediary services: Australia. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 51-66.
26. Li, N. and Krever, R., 2017. Loan intermediary services: China. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 109-122.
27. Lockwood, B. and Yerushalmi, E., 2019. How should payment services be taxed? *Social Choice and Welfare*, 53(1), pp. 21-47. https://doi.org/10.1007/s00355-019-01178-6
28. López-Laborda, J. and Peña, G., 2018. A new method for applying VAT to financial services. *National Tax Journal,* 71(1), pp. 155-182. https://doi.org/10.17310/ntj.2018.1.05
29. Mirrlees, J. [et al.], 2011. *Tax by Design: The Mirrlees Review*. Oxford: Oxford University Press.
30. Næss-Schmidt, H. S., Heebøll, C. and Lund, C., 2016. *Why and how to apply a Value Added Tax on financial services.* Copenhagen Economics.
31. Pallot, M. and Allen, T., 2017. Loan Intermediary Services: New Zealand. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 159-168.

32. Pardete, R., Santos, M. C. and Leote, F., 2024. Future VAT Regime for Financial Services from a Stakeholder Perspective: Analysis of the European Commission 2020 Public Consultation's Position Papers. *EC Tax Review*, 33(1), pp. 33-46. https://doi.org/10.54648/ECTA2024004
33. Peña, G., 2018. *Análisis económico de la tributación de los servicios financieros en el IVA*. Doctoral thesis. Universidad de Zaragoza.
34. Peña, G., 2026. Better VAT for Both Explicit and Implicit Financial Fees. *Public Budgeting and Finance.* forthcoming. https://doi.org/10.1111/pbaf.70021
35. Poddar, S. and English, M., 1997. Taxation of Financial Services under a Value-Added Tax: Applying the Cash-Flow Approach. *National Tax Journal*. 50(1), pp. 89-111. https://doi.org/10.1086/NTJ41789244
36. Slemrod, J. and Velayudhan, T., 2022. The VAT at 100: A retrospective survey and agenda for future research. *Public Finance Review,* 50(1), pp. 4-32. https://doi.org/10.1177/10911421221094595
37. Teijeiro, G., 2017. Loan Intermediary Services: Argentina. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 35-49.
38. Thang, S., 2017. Loan intermediary services: Canada. In: R. F. van Brederode and R. Krever, eds. *VAT and Financial Services: Comparative Law and Economic Perspectives*. Singapore: Springer Singapore, pp. 87-108.
39. Weisbach, D. A., 2019. Financial Services Under a Consumption Tax: Grubert and Mackie (2000) Revisited. *National Tax Journal*, 72(1), pp. 237-249. https://doi.org/10.17310/ntj.2019.1.09
40. WGS, 2018. *Taxing financial services, Discussion Paper for Session 11 of the Tax Working Group*.
41. Zee, H. H., 2005. A New Approach to Taxing Financial Intermediation Services under a Value-Added Tax. *National Tax Journal,* 58(1), pp. 77-92. https://doi.org/10.17310/ntj.2005.1.04

PUBLIC SECTOR ECONOMICS
50 (3) 431–460 (2026)

APPENDIX
MATHEMATICAL PROOFS

Theorem 1. Maximisation of profits in a business applying the MR method

$$\underset{\{L,D\}}{Max}\ \Pi = (\rho_L - \varepsilon_k)L + (\varepsilon_k - \rho_D)D \quad s.t.: \rho_L L + \rho_D D = TVI \Rightarrow \begin{cases} \rho_k := -\lambda = \dfrac{\rho_L - \rho_D}{\rho_L + \rho_D} \\ \varepsilon_k = \dfrac{2\rho_L \rho_D}{\rho_L + \rho_D} \end{cases} \tag{A1}$$

Proof: From the maximisation problem, using L and D as target and decision variables and TVI as the Total Value of Interest according to López-Laborda and Peña (2018), the Lagrangian function can be written:

$$l = (\rho_L - \varepsilon_k)L + (\varepsilon_k - \rho_D)D + \lambda(\rho_L L + \rho_D D - TVI) \tag{A2}$$

The First Order Conditions (FOC) are:

$$\begin{cases} \dfrac{\partial l}{\partial L} = (\rho_L - \varepsilon_k) + \lambda(\rho_L) = 0 \Rightarrow -\lambda\rho_L = \rho_L - \varepsilon_k \\ \dfrac{\partial l}{\partial D} = (\varepsilon_k - \rho_D) + \lambda(\rho_D) = 0 \Rightarrow -\lambda\rho_D = \varepsilon_k - \rho_D \end{cases} \Rightarrow \begin{cases} \rho_k := -\lambda = \dfrac{\rho_L - \rho_D}{\rho_L + \rho_D} \\ \varepsilon_k = \dfrac{2\rho_L \rho_D}{\rho_L + \rho_D} \end{cases} \tag{A3}$$

Q.E.D. ■

Theorem 2. Maximisation of profits in a business applying the Hybrid System

$$\begin{aligned} Max\ \Pi_L &= (\rho_L - \varepsilon_j)L \\ s.t.: \rho_L L &= IR \end{aligned} \Rightarrow \rho_{j,L} := -\lambda_L = \frac{\rho_L - \varepsilon_j}{\rho_L}$$

$$\begin{aligned} Max\ \Pi_D &= (\varepsilon_j - \rho_D)D \\ s.t.: \rho_D D &= IP \end{aligned} \Rightarrow \rho_{j,D} := -\lambda_D = \frac{\varepsilon_j - \rho_D}{\rho_D} \tag{A4}$$

Proof: The Lagrangian function can be written:

$$\begin{aligned} l_L &= (\rho_L - \varepsilon_j)L + \lambda(\rho_L L - IR) \\ l_D &= (\varepsilon_j - \rho_D)D + \lambda(\rho_D D - IP) \end{aligned} \tag{A5}$$

The FOC are:

$$\begin{cases} \dfrac{\partial l_L}{\partial L} = (\rho_L - \varepsilon_j) + \lambda_L(\rho_L) = 0 \Rightarrow -\lambda_L\rho_L = \rho_L - \varepsilon \Rightarrow \rho_{j,L} := -\lambda_L = \dfrac{\rho_L - \varepsilon_j}{\rho_L} \\ \dfrac{\partial l_D}{\partial D} = (\varepsilon_j - \rho_D) + \lambda_D(\rho_D) = 0 \Rightarrow -\lambda_D\rho_{iD} = \varepsilon - \rho_{iD} \Rightarrow \rho_{j,D} := -\lambda_D = \dfrac{\varepsilon_j - \rho_D}{\rho_D} \end{cases} \tag{A6}$$

Q.E.D. ■

PUBLIC SECTOR ECONOMICS
50 (3) 431-460 (2026)

Theorem 3. Maximisation of profits in a business applying the G method

$$\begin{array}{l} Max\ \Pi_i = (\rho_{iL} - \varepsilon_i)L + (\varepsilon_i - \rho_{iD})D \\ s.t.: \rho_{iL}L + \rho_{iD}D = TVI_i \end{array} \Rightarrow \begin{cases} \rho_i := -\lambda_i = \dfrac{\rho_{iL} - \rho_{iD}}{\rho_{iL} + \rho_{iD}} \\ \varepsilon_i = \dfrac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} \end{cases} \tag{A7}$$

Proof: The Lagrangian function is expressed as:

$$l_i = (\rho_{iL} - \varepsilon_i)L + (\varepsilon_i - \rho_{iD})D + \lambda_i(\rho_{iL}L + \rho_{iD}D - TVI) \tag{A8}$$

The FOC are obtained:

$$\begin{cases} \dfrac{\partial l_i}{\partial L} = (\rho_{iL} - \varepsilon_i) + \lambda_i(\rho_{iL}) = 0 \Rightarrow -\lambda_i\rho_{iL} = \rho_{iL} - \varepsilon_i \\ \dfrac{\partial l_i}{\partial D} = (\varepsilon_i - \rho_{iD}) + \lambda_i(\rho_{iD}) = 0 \Rightarrow -\lambda_i\rho_{iD} = \varepsilon_i - \rho_{iD} \end{cases} \Rightarrow \begin{cases} \rho_i := -\lambda_i = \dfrac{\rho_{iL} - \rho_{iD}}{\rho_{iL} + \rho_{iD}} \\ \varepsilon_i = \dfrac{2\rho_{iL}\rho_{iD}}{\rho_{iL} + \rho_{iD}} \end{cases} \tag{A9}$$

Q.E.D. ■

Institute of
Public Finance