Lund University
Department of Political Science

STVM23
Spring term 2026
Supervisor Maria Strömvik

Cash for reform

Did conditional EU funding deliver structural reforms in Member States?

Master thesis

Anna Selma Wentzer

Abstract

The EU's Recovery and Resilience Facility (RRF), a crisis response of €673 billion to the COVID-19 pandemic, made it possible to reward reforms in Member States for the first time. This thesis investigates whether the RRF can promote structural reforms, in line with the European Semester's country-specific recommendations (CSRs), using a principal-agent framework to theorise the RRF's novel structures of incentive and conditionality. The research design matches fulfilled RRF reforms to the 2019 and 2020 CSRs for all 27 Member States and analyses the relationship between each Member State's score of addressed CSRs and the size of its RRF envelope as a share of GDP. The analysis finds a positive and statistically significant relationship, robust after controlling for the number of CSRs and economic vulnerability. The incentive structure proved most effective for CSRs in energy, transport, and education policy, whilst the limits of the incentive structure are most visible in CSRs on pensions, taxation, housing, and business environment, areas characterised by high political salience and strong national competence. The answer to the research question is a cautious yes: to some extent, the EU can pay its Member States to undertake structural reforms.

Keywords: Recovery and Resilience Facility, Country-specific recommendations, Conditionality, Principal-agent problem, Reforms

Word count: 19.690

Reference system: APA Style 7th edition

Contents

Tables and Figures

1. Introduction

In summer 2020, the EU's Member States were looking down the barrel of a potentially devastating economic crisis in the wake of the COVID-19 pandemic. But as it turned out, the EU took the advice from the Global Financial Crisis of the chief of staff to US President Obama to heart: "*You never want a serious crisis to go to waste. (…) it's an opportunity to do things that you think you could not do before*" (Friedman, 2008, p. 1). Thus, its Member States' leaders agreed to form the Recovery and Resilience Facility (RRF) of 673 billion euros, funded by common debt issued by the Commission on behalf of the EU. Here is how it works: the RRF is supporting objectives for economic recovery in the EU's Member States through national investment plans (RRF plans) for the green transition, digitalisation, skills and education, and recovery of healthcare systems and businesses. The innovation to the functioning of the instrument is the feature of 'financing not linked to costs', which means the payments to Member States are released upon completing set objectives in their RRF plans, rather than documenting eligible costs. Hereby, the RRF payment model is performance-based instead of cost-based, which makes it possible to reward new sorts of activities – such as reforms (Zeitlin et al., 2025, p. 865).

Making structural reforms eligible for EU funding is a key novelty of the RRF and introduces two things the EU has 'not done before', or are at least relatively new: firstly, the financial incentive for Member States to commit to structural reforms during the planning phase of the RRF plans is brand new. According to the RRF regulation (2021/241, article 18), the reforms should address the EU's recommendations for structural reforms to improve economic conditions, called the country-specific recommendations (CSRs). In the past, implementing the CSRs was voluntary (unless part of macroeconomic surveillance after the Global Financial Crisis), with modest interest in their implementation in the Member States (Deroose & Griesse, 2014; Efstathiou & Wolff, 2018; Varga & Veld, 2013). While planning the reforms and their relation to the CSRs as part of the RRF plans largely remains at the discretion of the Member States, the RRF instrument provides a benefit for doing so.

Secondly, the Member State can only access the financial reward for implementing a reform if the Commission deems the fulfilment of the planned reform satisfactory, which means the instrument operates by conditional payments. Conditionality – the practice of attaching structural reform requirements to grants or loans from international institutions – is also a relatively new instrument in the EU programmes' toolbox. The international community has plenty of experience with conditional bailouts, investments, loans or partnership agreements; most of the time, these measures to implement reforms in the target country are considered a failure (Dreher, 2009; Efstathiou & Wolff, 2018; Khemani, 2017). The EU has previously and quite successfully employed conditionality in the accession process for the EU's candidate countries (Jacoby & Hopkin, 2019), but not in relation to the EU budget before 2021 (Zeitlin et al., 2025, pp. 864–865).

While the sheer size of the RRF or the decision to use incentives and conditionality to reward structural reforms warrants scholarly attention, the implications of this development are magnified by the new proposal for the next EU budget from 2028 to 2034. According to the Commission's proposal, almost half of the budget, roughly 865 billion euros, should be paid out through national investment plans following the performance-based model. Like in the RRF regulation, the plans can include reforms, and part of the guidance for the Member States to plan those reforms will be the CSRs (Commission, 2025, art. 22).

If the RRF's model for promoting structural reforms by financial incentives acts as a blueprint for the EU's future budget, the question arises as to what the experience with the RRF in this respect has been. Previous studies of the RRF have focused on economic effects, European integration or reforms across policy areas but not related to the CSRs (Guidi et al., 2023; Kalyuzhna & Shchelkunov, 2025). Furthermore, the Commission evaluates the implementation rate of the CSRs, but without a dedicated focus on the impact of the RRF (European Commission, n.d.). The results of rewarding Member States to implement structural reforms remain unknown, and I therefore intend to answer the following research question:

Can the EU promote structural reforms in its Member States by rewarding them financially?

This thesis will answer this research question by analysing the uptake of the EU's recommendations for structural reforms, the country-specific recommendations (CSRs), through the incentive and conditionality structure that promotes reforms in the Recovery and Resilience Facility (RRF) – the measure of the financial incentive is the size of the Member States' envelopes.

The RRF is designed in keeping with new public management principles – the idea that policy outcomes can be measured, benchmarked, and incentivised through a performance-based payment model (Thinus, 2025, p. 883). The RRF's payment model, based on national RRF plans, distinguishes it from the externally imposed conditionality that the broader literature has consistently found ineffective, because the Member States determine what their payments should be conditional on (Arnal et al., 2025, p. 78). This output-oriented logic fits the principal-agent framework as developed by Jensen and Meckling (1976), and provides the underlying assumption of this thesis: that a financial incentive can influence national reform policy. The framework is well-suited to the two-phased governance architecture of the RRF: the planning phase of the RRF plans, which ends when they are adopted by the Council of the EU, and the implementation phase, where payments are made conditional upon the fulfilment of investments or reforms detailed in the RRF plan. Whether this design succeeds in producing alignment between the individual Member State's behaviour and the EU's and European Member States' general reform ambitions is the empirical question this study addresses.

The research design consists of two sequential steps. The first step matches 1,617 activities across 912 fulfilled reforms to the country-specific recommendations (CSRs) from 2019 and 2020 for all 27 Member States, producing a score for the share of addressed CSRs for each Member State and each of 17 policy categories. The second step finds that there is a moderately positive relationship between the scores for the Member States and the size of each Member State's RRF envelope as a share of GDP, relying on a scatterplot, Pearson's r, Spearman's ρ, and OLS regression. The study treats the financial incentive as an exogenous explanatory variable and focuses exclusively on implemented and financially rewarded reforms, as they are the outcome of both the incentive and conditionality structures through planning and implementing the RRF plans. It does not assess endogenous domestic factors, the macroeconomic effects of the RRF, planned but unfulfilled reforms, or the depth of individual reform alignment – limitations that are elaborated in the following sections.

The thesis contains 7 sections, which proceed as follows. Section 2 takes stock of the current literature by tracing three intersecting bodies of research on understanding policy change, the barriers to reform, and the experiences with conditionality before situating the RRF within that literature as a distinct governance instrument. Section 3 elaborates on how the principal-agent framework can be used to understand the RRF's two-phase governance structure and the limits of the instrument. Section 4 details the methodological considerations, the research design, and the variables for testing the relationship between the score of addressed CSRs and the size of the RRF envelope for each Member State. Section 5 presents the findings: there is a moderately positive relationship between the uptake of CSRs and the size of the RRF envelope, the Member States preferred policy areas with existing EU enforcement, and Hungary is a structural outlier, excluded from the study of correlation. Section 6 discusses the findings in relation to the research question and relevant literature. To conclude, section 7 outlines how the findings provide a cautious affirmative answer to the research question and what further lessons can be learned from these findings.

2. Reform, conditionality, and the European Semester

The Recovery and Resilience Facility represents a significant evolution in how the European Union encourages national reform efforts. Understanding this development requires tracing three intersecting bodies of literature: what constitutes a reform and how policy change occurs; why economically beneficial reforms are frequently blocked by political constraints; and how conditionality as an instrument has been used – and largely failed – both by international financial institutions and, more recently, by the EU itself. The first strand establishes that a reform is a contested concept, with scholars diverging on questions of scope, intent, and measurement, and that theories of policy change struggle to be applicable ex-ante. The second strand, drawing on the political economy of reform literature, shows that even where the case for reform is clear, domestic barriers rooted in credible commitment problems, norms, and preferences for public goods can prevent reforms from taking hold – with particular implications for externally imposed reform efforts. The third strand traces how conditionality has evolved within the EU from a lever facilitating accession to a club imposing macroeconomic discipline, and how the European Semester's country-specific recommendations (CSRs) have struggled to drive structural reform in the absence of credible financial incentives. Together, these strands reveal a fundamental tension between the EU's ambition to steer structural reform and the institutional and political constraints that limit its effectiveness – a tension the RRF attempts to overcome, and that this thesis examines empirically.

2.1 Reform and policy change: what is it and how is it studied?

The concept of reform lacks a commonly agreed-upon definition in the political science literature, a condition that complicates both comparative analysis and policy evaluation. Several scholars flag this lack of a common conceptualisation and definition as a challenge (Basil & Nwankwo, 2010, p. 186; Capano, 2009, p. 8; Cerna, 2013, p. 4). Generally, scholars from political science and political economy agree that a reform constitutes deliberately pursued change within an existing system – as opposed to revolutionary rupture – but diverge significantly on questions of scope, intent, and measurement (Capano & Howlett, 2009, p. 2; Khemani, 2017). A frequently cited illustration of this complexity is Peter Hall's (1993) three-orders framework, which distinguishes routine instrument adjustments, non-routine changes of instruments within a stable policy goal, and wholesale paradigm shifts following a crisis. This framework effectively introduces levels of reform or policy change derived from their core functions in relation to the existing system. A complementary challenge is raised by Khemani (2017, p. 15), who argues that only changes producing a better economic outcome than the status quo ought to be counted as reforms – introducing the further problem of how and when to measure effects, another challenge in the broader literature (Capano, 2009, p. 9; Capano & Howlett, 2009, p. 2; Cerna, 2013).

But how to study reforms or policy change? Streeck and Thelen (2005) reject the widespread notion of assessing policy change based on a punctuated equilibrium model and develop a framework by accounting for types of processes and results of change (Streeck & Thelen, 2005, p. 9). Capano and Howlett (2009) argue that in the face of ambiguous concepts, the epistemological and methodological premises for studying policy change are crucial. They derive four theoretical distinctions of change from the body of literature: the cyclical, the dialectic, the linear, and the teleological. That theoretical starting point is the basis for whether the emphasis should be on structure (cyclical, teleological) over agency (linear, dialectic), endogenous (dialectic) or exogenous factors or focus on output (linear, teleological) rather than process (Capano & Howlett, 2009, pp. 2–3). Studying reforms or policy change remains a question of methodology and epistemology, as definitions of the process and the result depend on what the analysis emphasises. A different approach is to study why a reform, or policy change, does *not* take place – this is a feature in the political economy of reform, rooted in the understanding of reforms as structural reforms leveraging broad economic prosperity.

The literature of political economy of reform examines why economically beneficial reforms (structural reforms) are frequently blocked, focusing on political incentives, power distributions, and institutional constraints rather than the technical merits of reform proposals. Fernandez and Rodrik (1991, p. 1146) find that a barrier to structural reforms is the inability to determine the beneficiaries of reforms ex ante; this substantially alters the rationale for the political agents. Alesina and Drazen (1991, p. 1171) theorize that needed reforms are dragged out until the economic situation is dire, because if the cost of the reform is asymmetric across groups there is an incentive to wait each other out; subsequently, this is the reason for many structural reforms being implemented in times of crisis. Khemani (2017) identifies three distinct strands of explanation for political constraints to reform: problems of credible commitment, problems of norms, and problems of preferences for public goods. Acemoglu (2003) illustrates the credible commitment problem: groups cannot resolve distributional conflicts over reforms through negotiation alone, because no credible mechanism exists to prevent politically powerful actors from reneging on any agreement once reached. Problems of norms can be rooted in non-cooperative norms or disagreements about the nature of the problem; this in turn, leads to different expectations towards the role of government and thus preferences for public goods(Khemani, 2017, p. 3). These barriers to structural reforms create the very opening that conditionality as an external instrument is designed to exploit – and, as the following section shows, frequently fails to fill.

2.3 Forcing reform: conditionality

The use of conditional financial assistance as a tool for driving structural reform is widely considered ineffective, most extensively documented in the context of the International Monetary Fund (IMF) and the World Bank. Historically, conditionality consists of requirements that the debtor meets specific conditions on macroeconomic policy when international organisations, the creditors, provide loans (Bekker, 2021, p. 179). Dreher (2009) defines the aim of conditionality in an IMF context as improving the economic outlook of target countries through institutional change, yet finds that recipient countries overwhelmingly continue to rely on IMF support over time, suggesting that the conditions attached to loans fail to produce the intended self-sustaining improvements. This ineffectiveness has been attributed to a range of structural barriers: Feldstein (1998), Radelet and Sachs (1998, p. 4), and a Congress-appointed Commission (Meltzer, 2005, p. 3) point out how the IMF is simply lacking the expertise and legitimacy of imposing structural conditionalities, drawing away political capacities in the target countries from domestic problem-solving and towards satisfying external demands that may be ill-fitted to specific national contexts. This legacy makes it all the more striking that the European Central Bank, Commission, and Council resorted to conditional financial assistance when responding to struggling Member States in the aftermath of the Global Financial Crisis — a choice whose origins and consequences are examined in the following section (Jacoby & Hopkin, 2019).

Despite the well-documented failings of conditionality in IMF and World Bank settings, the EU resorted to conditional financial assistance in the aftermath of the Global Financial Crisis, a choice that Jacoby and Hopkin (2019) argue is best explained by the EU's own prior experience with the instrument. They find that the European institutions all came to view conditionality as a promising tool on the basis of its apparently successful application during the EU enlargement process of the early 2000s, when candidate states were required to meet reform conditions prior to accession. Jacoby and Hopkin (2019, p. 21) characterise this earlier use as a 'lever' – conditionality that assisted reform-minded national elites towards a clearly defined and credible reward, namely EU membership. For contrast, conditionality was used as a 'club' in the Eurozone debt crisis, where financially distressed Member States were subject to extensive and intrusive conditions on core state functions, including fiscal, labour market, and social policy (Jacoby & Hopkin, 2019, p. 24). Also during the Eurozone debt crisis, the EU developed a framework for coordinating Member States' economic policies: the European Semester, which combines the Stability and Growth Pact, the Macro-Economic Imbalance Procedure, and social and employment coordination into a single annual cycle, producing country-specific recommendations for each Member State (Bekker, 2021, p. 176; Efstathiou & Wolff, 2018, p. 2). However, as the following section discusses, the Member States have struggled to implement the structural reforms encouraged in the recommendations.

2.4 The European Semester and the Recovery and Resilience Facility

The country-specific recommendations (CSRs) produced by the European Semester have consistently fallen short of their intended effect, with implementation rates insufficient to generate the coordinated economic benefits they were designed to produce. Four years after the CSRs were first introduced, Deroose and Griesse (2014, p. 1) found an implementation rate of approximately 40 pct – higher than critics had anticipated, but not sufficient to unlock the growth and employment benefits associated with coordinated structural reform across Member States (Varga & Veld, 2013). The level of implementation was found to vary with the electoral cycle and market pressure in the member state concerned, and to increase where the EU held meaningful enforcement powers – suggesting that voluntary compliance with recommendations is structurally limited (Deroose & Griesse, 2014; Varga & Veld, 2013). Several years later, Efstathiou and Wolff (2018, p. 15) find that implementation rates had slowed further and identify a fundamental dilemma at the heart of the Semester's design. European policy coordination is justified because national policies on pensions, labour markets, or childcare affect the functioning of the Internal Market, yet national governments and parliaments retain competence by virtue of the division of competences and the principle of subsidiarity. In the framework of Jacoby and Hopkin (2019), the European Semester is equipped with neither a lever nor a club – it can neither offer a sufficiently credible reward for compliance nor compel it. The introduction of the Recovery and Resilience Facility (RRF) represented an attempt to address this gap by attaching conditional financial support to the CSRs, as discussed in the following section.

Looking at the CSRs, the introduction of the RRF marked a significant shift in the European Semester's capacity to drive structural reform, as the EU could reward structural reforms financially, creating an incentive for Member States to follow recommendations. Prior attempts to strengthen the Semester's effectiveness – including changes introduced in 2015 to increase the involvement of national actors and improve national ownership – had failed to produce more ambitious uptake of the CSRs (Bekker, 2021, p. 176; Efstathiou & Wolff, 2018, p. 15). The RRF, established in February 2021, offered €672.5 billion in loans and grants to Member States on the condition that their national investment plans (RRF plans) be consistent with, amongst other requirements, the CSRs issued under the Semester (Zeitlin et al., 2025). As Bekker (2021, p. 180) sets out, this means that the CSRs became one of the conditions for receiving financial support and fed directly into the criteria against which the use of grants and loans would be evaluated, giving the CSRs a financial incentive for compliance that they had previously lacked outside the macroeconomic imbalance procedure. Whether this design succeeds where the earlier voluntary and consensus-based efforts failed depends on the criteria for success determined in the negotiations.

The conditionality embedded in the RRF was not introduced solely to strengthen the effectiveness of the European Semester but was also a political concession required to secure agreement on the unprecedented common debt instrument underpinning the fund. Thinus (2025, p. 883) argues that the conditions attached to the RRF are a result of economic neoliberalism and instrumentalise the CSRs as a way to protect European values and financial interests. The financing of the RRF through common European debt was met with significant resistance from the Netherlands, Denmark, Sweden, and Austria, collectively sceptical of debt mutualisation and the redistributive implications of a fund whose grants would flow disproportionately to Member States in Southern and Eastern Europe (Bekker, 2021, p. 180). De la Porte and Jensen (2021, p. 2) find that strict conditionality on the rule of law and the CSRs was central to ensuring that these creditor Member States had sufficient oversight and control. This interpretation is consistent with Jacoby and Hopkins' (2019) observation that conditionality in the Eurozone debt crisis served partly to signal to Northern European creditor States that financial assistance would not reward profligacy – a political function distinct from, and at times in tension with, its reform-driving function.

The existing literature has not yet produced a systematic analysis of the nature of reforms implemented under the RRF relative to the CSRs issued under the European Semester. Whilst several studies have examined the distribution of RRF funding across policy areas and its projected macroeconomic effects, these analyses have largely been conducted based on Member States' planned investments and reforms rather than implemented ones (Bekker, 2021; Guidi et al., 2023). The European Commission monitors the implementation of the CSRs, but not limited to the contribution of the RRF; a study by the European Court of Auditors on the RRF’s engagement with CSRs on the business environment finds that much of the effect measured by the

Commission is happening in parallel with the RRF (ECA, 2025b). Guidi, Piccinetti and Verzichelli (2023, p. 346) compare Member States' initial national investment plans (RRF plans) to the policy fields of the European Semester but explicitly refrain from comparing the RRF plans directly to the CSRs, citing methodological obstacles including the difficulty of coding heterogeneous material across Member States and policy areas. As the RRF enters its final semester of implementation, these obstacles are substantially reduced: standardised reporting by Member States on implemented reforms now provides a common basis for systematic comparison that was unavailable to earlier studies, making a direct assessment of implemented reforms against the CSRs more feasible.

The RRF represents a distinct attempt to construct a conditionality governance system that departs from the patterns documented in the broader literature. The studies of the IMF and World Bank’s use of conditionality identify a persistent failure rooted in externally imposed conditions that lack domestic legitimacy and are applied to sovereign states by institutions without the expertise or political standing to enforce structural change (Dreher, 2009; Feldstein, 1998; Meltzer, 2005, p. 3). The RRF, by contrast, is built on an ex ante reward structure in which Member States define their own reform commitments in advance and receive financial rewards only upon documented fulfilment – closer to the lever dynamic that Jacoby and Hopkin (2019, p. 24) identify as the condition under which conditionality is more likely to succeed. This reward structure is, as Thinus (2025) argues, a product of new public management principles: the idea that policy outcomes can be measured, benchmarked, and incentivised. It is precisely this governance architecture – The EU offering a pre-defined financial reward to a Member State in exchange for reform commitments, with conditionality enforced against the Member State's own RRF plan – that motivates the principal-agent framework as the theoretical lens of this study, as elaborated in the following section.

3. Incentives and conditionality: a two-phase governance logic

The incentive structure of the Recovery and Resilience Facility (RRF) can be understood as two sequential phases: the planning phase, where Member States plan their national investment plan (RRF plan) in dialogue with the Commission, and the implementation phase, where the Member States document their implementation of the RRF plan to the Commission. The RRF is a novel and temporary instrument, concluding in December 2026, and its institutional setup is complicated by the dual role of Member States. Member States have collectively agreed on the instrument, participated in producing their own CSRs, and are largely in the driver's seat when designing their RRF plans. At the same time, they are individually subject to the enforcement of those same plans. The first phase is the planning phase. Here, the intergovernmental level – the European Council and the Council of the EU – functions as the principal. It offers Member States a financial incentive: access to a dedicated and tailored envelope of the RRF in exchange for committing to an RRF plan, including reforms consistent with the CSRs. The second phase begins once the plans are adopted by the Council. The principal shifts to the Commission, which makes access to the RRF envelope conditional on each Member State's RRF plan. Payments are made only in return for documented results – completed investments and reforms. This shift is the defining structural feature of the RRF and creates a tension between the European-level interest in shared economic stabilisation and the domestic interest in minimising the political costs of reform – a tension the principal-agent framework is well suited to analyse.

The principal-agent framework, as developed by Jensen and Meckling (1976), sets out how divergent interests generate predictable incentive problems in the governance relationship at the heart of the RRF. An agency relationship is defined as follows:
"(…) a contract under which one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision-making authority to the agent. If both parties are utility maximisers, there is good reason to believe that the agent will not always act in the best interests of the principal" (Jensen & Meckling, 1976, p. 308)
To limit divergence from its interests, the principal can instate incentives and incur 'monitoring costs'. In turn, the agent carries 'bonding costs' to reassure the principal of its intentions – for example, through reporting for transparency purposes. Jensen and Meckling (1976, p. 308) argue that some divergence will always persist, resulting in a 'residual loss' for the principal. A core assumption of this thesis is that all actors are rational and act in a utility-maximising manner. A financial incentive, therefore, has direct explanatory value for the agent's behaviour. This framework has previously been applied to the relationship between voters and politicians (Ferejohn, 1986, p. 8), and between European governments and the Commission and European Court of Justice (Hooghe & Marks, 2001, p. 10).

Another concept within the principal-agent relationship that is particularly relevant to the RRF's governance design is the ‘moral hazard problem’, which is connected to the residual loss. The moral hazard problem describes what follows from the agent's incentive to act in its own interest rather than the principal's, exploiting the fact that its actions cannot be fully observed or verified (Busato & Coletta, 2017, p. 299). In the context of the RRF, this risk is most pronounced in the implementation phase, where the Commission relies on the Member State's own documentation to verify compliance, and the European Court of Auditors has criticised this documentation as occasionally weak and heterogeneous across Member States (ECA, 2024, pp. 396–370). Because the reward is tied to documentation rather than to an independent assessment of reform quality, the moral hazard problem places a structural limit on what the monitoring architecture can achieve – meaning the residual loss for the principal may reflect not only domestic political resistance to reform, but also performative or superficial implementation. The following sections apply the framework to the two phases of the RRF identified above.

3.1 The planning phase: funding as an incentive to reforms

At the planning phase, the intergovernmental level – made up of the European Council and the Council of the EU – functions as the principal by establishing the collective goal of structural reform and economic recovery and offering financial incentives to Member States to commit to it. The European Council agreed on the debt-financed instrument and its architecture to ensure economic recovery after the pandemic and to make funding available to reward reforms in line with the CSRs (Armingeon et al., 2022, pp. 160–161). The CSRs themselves are adopted by the Council of the EU after discussion in the European Council. According to Varga and Veld (2013), Member States share a common interest in national structural reforms, as these can leverage growth and employment across the single market. The combined investments and reforms in the RRF plan should have an estimated cost corresponding to the Member State’s envelope (ECA, 2025a, p. 18) – but the payments on completion of individual RRF reforms do not need to reflect a cost, making the financial incentive attractive regardless of reform difficulty. This incentive structure is the mechanism through which the principal sought to align Member States' planning choices with the CSRs, as examined at the national level in the following section.

The national level is where the principal's financial incentive meets domestic politics in the planning phase – and where the RRF may address one of the most persistent barriers to structural reform. The political economy of reform literature identifies asymmetric short-term costs as a key obstacle. Even when the long-term gains from a policy change are broadly distributed, the immediate costs fall on specific groups – such as workers facing restructuring or pensioners facing adjusted entitlements. This makes political opposition disproportionately strong relative to diffuse benefits (Alesina & Drazen, 1991; Fernandez & Rodrik, 1991; Khemani, 2017). The RRF's funding structure may address this barrier directly. By making EU grants available upon completion of a planned reform, the national government acquires resources that can be deployed to soften the financial impact on those bearing the short-term costs. A tax benefit for more sustainable transport vehicles might be blocked by having to increase revenue or cut spending, but it becomes more tractable if the cost can be funded through RRF payments. This logic is consistent with Jacoby and Hopkins' (2019) finding that conditionality works best when it assists reform-minded national elites towards a credible reward. However, the concept of the residual loss predicts that the incentive structure will not ensure full implementation of the CSRs through the RRF reforms – some reforms will be subject to too much domestic resistance to be resolved by a check.

The planning phase was also shaped by significant time pressure, which could have interacted with the asymmetric incentive structure to produce different negotiation dynamics across Member States. The RRF regulation was adopted in February 2021, and Member States were eligible for pre-financing of up to 13 pct. of their envelope upon Council adoption of their RRF-plan, provided this occurred before 31 December 2021 – leaving less than a year to design, negotiate, and submit a national investment plan of unprecedented scope (Regulation (EU) 2021/241, art. 13). In line with the principal-agent framework, this deadline created an asymmetric lever: for Member States with larger envelopes relative to GDP, the pre-financing represented a substantial financial reward that intensified the incentive to reach agreement on a plan with the Commission quickly. Meanwhile, Member States with smaller envelopes faced less pressure and had to negotiate a smaller plan. However, Bekker (2021, p. 182) highlights the notion in the regulation stating that the Commission should fully respect national ownership, while Eihmanis (2025) finds that the Commission undoubtedly shaped Estonia's RRF plan during negotiations. The effect of time pressure on the effect of the incentive is excluded from this study, as it is difficult to assess the effect on the planned reforms in the RRF plans. Therefore, this thesis relies exclusively on implemented and financially rewarded reforms rather than planned ones.

3.2 The implementation phase: conditionality on Member States

Once the RRF's national investment plans are adopted, the Commission assumes the role of principal, enforcing the financial conditionalities each Member State has itself agreed to by withholding payments unless documented results are delivered. Under the RRF's performance-based payment model, funding tranches are released only upon the Commission's verification that agreed milestones and targets have been met. If a Member State wishes to amend its planned reforms, this requires approval from both the Commission and the Council. In principal-agent terms, the Commission's verification and payment procedures constitute the principal's monitoring of expenditures. However, the Commission's monitoring capacity faces a structural limitation. There is no legal definition of what constitutes a reform – mirroring the challenge flagged in the political science literature (Cerna, 2013, p. 4). The milestones and targets against which compliance is verified were negotiated with the Member States in their RRF plans, and the Commission relies on the reported documentation. Throughout the RRF implementation, the controls and reporting frameworks have been updated – for instance, if previously fulfilled milestones and targets were not fulfilled anymore, the Commission can subtract funding from the next payment to that Member State (ECA, 2025a, p. 32). The weaknesses of the monitoring instrument mean that residual loss may take the form of Member States mimicking compliance with the CSRs or their RRF plans rather than delivering the reforms the principal intended.

In many Member States, a new government had to implement reforms that were entrenched in the RRF plan negotiated by their predecessors to access their RRF envelope. While the Member States could choose to forego the funds and drop any loans they have committed to, the billions of RRF grants will be repaid by the EU budget until 2058; the Member States will thus participate in the repayment of the RRF grants regardless of their own benefit (Arnal et al., 2025). Member States can make amendments to their RRF plans, but only due to "objective circumstances" as stated in the regulation (2021/241, art. 21), which then go through adoption by the

Council. Examples of this struggle are the cases of Italy, having to politically compromise during the implementation, and Sweden, engaging in lengthy renegotiations of its RRF plan. In Italy, the RRF plan was negotiated by Mario Draghi's government, but Giorgia Meloni was responsible for the implementation from 2022 onwards and attempted a complete overhaul, ultimately having to continue with some adjustments to access the biggest envelope in the RRF (Capati & Christiansen, 2025). In Sweden, a new government took office shortly after the RRF plan was adopted in 2022, and in 2024, several amendments were adopted: the justifications were on the grounds of better methods and the rise of costs due to inflation, but the result was fewer administrative tasks and the removal of milestones for a fuel reduction obligation (Commission, 2024). Both effects align with a shift in political priorities between the two governments (Aylott & Bolin, 2023).

Compliance with national RRF plans against or at odds with the sitting government creates a different set of bonding costs for the agent and risk of further residual loss. The national ownership component of the RRF plans created in the planning phase might be challenged, jeopardising the alignment of interests between the Member State and the EU. This could fan the moral hazard problem, increasing superficial compliance to extract funding without real effort to implement the reforms and realise their objectives. Additionally, the Member State could decide to simply drop the implementation of a given planned reform, adding to the residual loss. While the principal-agent framework assumes that the agents' self-interest can be adjusted by economic gain, the political economy of reform literature adds the notion of ownership as an important factor to the uptake of conditional reforms. The governance structure of the RRF ensures significant ownership and extensive involvement by the Member States' governments, but the level of ownership can decrease with changes of government, adding to the risk of residual loss by not implementing planned reforms.

4. From recommendations to reforms: research design and variables

Having established uncharted territory in previous literature and the theoretical basis for prescribing an effect to financial incentives and conditionality, this section elaborates on the methodological considerations behind developing a fitting research design. The country-specific recommendations (CSRs) serve as both the legal and institutional focal point of this study, as they are anchored in the regulation establishing the Recovery and Resilience Facility (RRF) and define the direction of the principal-agent relationship theorised above. According to Capano and Howlett (2009, p. 2), a clear policy goal and external steering of this kind align with a teleological understanding of policy change, in which Member States are directed towards defined reform objectives and outcomes are evaluated against those pre-stated goals. This has a specific epistemological implication for the research design: the focus is on output rather than process, meaning that only reforms officially classified as fulfilled and resulting in payment are included in the analysis. Reforms that indicate intention but have not been verified and disbursed by the Commission are therefore excluded, since the purpose is to evaluate the results of the incentive structure as an instrument – what it delivers – rather than to reconstruct how individual reforms were planned or negotiated domestically.

This research treats the RRF's financial incentive as an exogenous explanatory factor, rather than seeking to account for the endogenous domestic conditions that have also driven reform efforts in individual Member States. This choice follows directly from the principal-agent framework: if financial incentives are the cross-cutting mechanism through which the principal seeks to align agent behaviour, then variation in the size of that incentive across Member States should produce variation in the uptake of the CSRs – making it the theoretically motivated independent variable. The aim is therefore to assess the effect of a single, asymmetric incentive structure across all 27 Member States, in line with Capano and Howlett's (2009, p. 2) observation that teleological models of policy change foreground exogenous variables over internal political dynamics. A descriptive analysis of reform patterns by policy area serves as a complement to this, mapping what the principal has gained and lost across Member States in aggregate – revealing which policy areas the incentive structure has been most and least effective at driving, and thereby giving more texture to the residual loss identified in the principal-agent framework.

The research design is deductive in its overall orientation, testing a theoretically derived expectation – that larger financial incentives produce higher CSR uptake. However, the first step of the analysis, which forms the empirical foundation for the subsequent correlation, required an iterative approach: the method for matching fulfilled reforms to the CSRs was developed with an empirical starting point, as no established coding framework exists for this novel area of analysis. This iterative necessity is compounded by a structural feature of the RRF regulation itself: there is no commonly agreed definition of what constitutes a reform. This reality is reflected in the heterogeneity of the empirical material across Member States: the reported reforms span everything from roadmaps, amendments, new regulations, programmes, task forces, and strategies. Attempting to categorise or value reform types would therefore introduce a layer of analytical judgement unsupported by either the regulatory framework or an established theoretical framework for categorising scope, intent, or measurement (Cerna, 2013, p. 4; Regulation (EU) 2021/241, 2021; Zeitlin et al., 2025). To ensure that interpretive decisions remained traceable and consistent across all 27 cases, I wrote memos throughout the matching process to document the basis for coding judgments, the criteria applied, and any revisions made, as elaborated in a dedicated section on RRF reforms as a variable.

4.1 Research design

To examine the relationship between the RRF's incentive structure and the uptake of the CSRs, the research design consists of two sequential steps: matching CSRs to fulfilled reforms across all 27 Member States and then analysing the correlation between the resulting CSR uptake scores and each Member State's financial allocation. The first step is a systematic content analysis across 27 cases, in which fulfilled reforms are coded against a bespoke coding scheme – one per Member State – constructed from that Member State's CSRs. The CSRs are labelled with 'policy areas' by the Commission, and the matching process therefore also reveals which policy areas have been addressed and which have not across all Member States. The second step analyses the relationship between the financial incentive, operationalised as the national RRF envelope, and the Member States' results from the first step. Besides bivariate statistical analysis, the relationship is controlled for divergence in the number of CSRs and economic vulnerability. The data collection relies on an unobtrusive method drawing exclusively on official sources published by the Commission (2021) under the transparency requirements of the RRF regulation (2021/241, art. 25), ensuring that the empirical basis is consistent and comparable across all 27 cases (Clark et al., 2021b, pp. 306–307).

A longitudinal study comparing CSR uptake before and after the introduction of the RRF is not a viable alternative design for two reasons. First, the nature of the CSRs – which are frequently rolled over to the following year if the underlying challenge persists (Arnal et al., 2025) – means that the pre- and post-RRF periods cannot be cleanly separated on the basis of the recommendations themselves; the Commission actively uses the RRF as a cut-off point in its tracking (European Commission, n.d.). Second, any pre-post comparison would face the additional confounding problem of disentangling the effect of the RRF from the effect of the COVID-19 pandemic itself, which independently altered both the content of the CSRs and the reform priorities of Member States. The cross-sectional design is therefore a deliberate methodological choice. The outcome of the analysis, a descriptive overview generated in the first step, is driven by the missing insights into the relationship between RRF reforms and the CSRs in general, and the ambition to reveal more concretely what the financial incentive has delivered, as determined in the second step, across which Member States and for which policy areas. The two steps of the research design and the variables underlying them are elaborated in turn in the following sections.

4.2 The first step – matching CSRs and RRF reforms

The first step consisted of matching each fulfilled reform to a corresponding CSR for each Member State, comparing 1,617 activities across 912 reforms to the CSRs from 2019 and 2020 – the two cycles that served as the basis for the planning of the RRF plans under the RRF regulation. The data draws on archive material published by the Commission in line with the transparency requirements of the RRF regulation and the process of the European Semester, which is the annual cycle through which the Commission monitors Member States' economic and fiscal policies and issues the CSRs (Varga & Veld, 2013). The CSRs are labelled with policy area designations by the Commission, and matched reforms are required to relate to the same policy area, which produced a score of addressed sub-level CSRs for each Member State and each policy area, respectively.

The criteria for the matching process were a fit between the content of the RRF reform and any given CSR for the same Member State. During the process, I labelled all reforms as fitting, not fitting, or undecided, with the undecided category introduced after the second case to capture borderline matches. After completing all 27 cases, the first three were reassessed to check whether coding judgements had shifted, resulting in one reform being reclassified. All undecided reforms were subsequently resolved by consulting the full CSR documents to better understand the macroeconomic context behind each recommendation, since these documents elaborate on Member States' macroeconomic situations and the structural challenges underpinning each recommendation. The fitting matches fell on a spectrum between apparent and latent, given the variation between the reforms implemented by the Member States: Some reforms related directly to concepts in a CSR, whilst others more indirectly pursued an objective laid out in a recommendation. An example of an apparent match is Italy's reform "National Plan tackling undeclared work", which was matched to the CSR "(…) Step up efforts to tackle undeclared work" (CSR Italy, 2019). An example of a more latent match is Spain's "Law on the Quality of Architecture and Building Environment and New National Architecture Strategy" which was matched to the CSR "[Focus investment-related economic policy on] (…) and energy efficiency, (…)" (CSR Spain, 2019), as the content of the reform and milestones and targets latently contributes to improve the energy efficiency of buildings. Each reform was compared to all the CSRs in the corresponding Member State to decide on a single CSR as the best match.

All the reforms have one or several CSRs prescribed by the RRF plans; I chose to only match the reforms to one single CSR using the principle of matching the primary aim of the content of the reform to the best-fitting recommendation. I decided not to limit the potential matches to the prescribed CSRs, because the Member States have an incentive to report reforms against as many CSRs as possible, since the RRF regulation (2021/241, art. 17 and 18) requires that plans address all or a significant subset of the recommendations. This behaviour is consistent with the moral hazard problem identified in the principal-agent framework: because the reward is tied to documentation rather than to an independent assessment of reform quality, the agent has a rational incentive to present measures as CSR-responsive even where the substantive alignment is weak. This process did not track the cases in which a reform was matched to a different CSR than the reported one, as the cases were relatively few, and it is unclear what the criteria for prescribing CSRs to reforms are for the Member State or the Commission.

A limitation of this process is that the RRF reforms are matched to the CSRs regardless of the effect or extent to which they address the relevant recommendation, and this study can therefore not measure the actual compliance rate of the RRF reforms with the CSRs. The dataset by the Commission detailing CSRs uses a system for labelling the degree of implementation for most recommendations: no progress, limited progress, some progress, substantial progress, and full implementation. However, these labels cannot be operationalised to assess the effect of the RRF reforms, as it takes all policy measures in the given Member State into account and not just the RRF reforms or investments. To get a brief assessment of any patterns between the Commission's labels and Member States' RRF reforms, I translated the labels into the values 0, 25, 50, 75, and 100, respectively, as done in the study by Efstathiou and Wolff (2018). Then I calculated the average value for each Member State and compared it to the score of addressed CSRs through the RRF; there was no apparent pattern. However, this circumstance can not conclude whether the RRF reforms have contributed to the implementation of the CSRs; rather, it illustrates how the challenges in the CSRs from 2019 and 2020 vary in difficulty across Member States. For instance, the recommendation to invest in clean and efficient production of energy, recommended to all Member States, is given to a highly diverse status quo for the domestic production of renewable energy across the EU (Wójtowicz, 2024, p. 162). The following sections elaborate on the operationalisation of CSRs and RRF reforms for the purpose of this study.

4.2.1 The country-specific recommendations (CSRs)

The European Semester is a yearly cycle, and the relevant cycles for the matching process are 2019 and 2020, as their CSRs are the common starting point for all the Member States' planning phase of the RRF in line with the regulation (2021/241, art. 17 (3)). Restricting the coding scheme to the 2019 and 2020 CSRs serves both the internal consistency of the score calculation and the cross-country comparability of the analysis. Later recommendations, such as the 2024 cycle, could be the basis for amendments to the RRF plans introducing new reforms, but including these reforms would inflate the denominator of the CSR uptake score for the affected

Member States, as then all further CSRs from, e.g., 2024 would also be counted as addressed or not addressed. Fifteen reforms across 9 Member States are linked to CSRs issued after 2020 and are therefore excluded from the analysis. An example is Denmark's reform to reduce emissions of greenhouse gases from production on carbon-rich soils, which did not match the 2019 or 2020 CSRs as it was added later based on the 2024 CSR: "*take further efforts for sustainable agriculture by stepping up decarbonisation measures and action to reduce nutrient losses*" (CSR Denmark, 2024). Had the CSR been included in the study, Denmark would have gone from a total of 5 to 8 CSRs. Had all the 15 excluded reforms been included, those 9 Member States would have had more CSRs to cover, and the cross-country comparison would be methodologically challenged.

To convert the CSRs from 2019 and 2020 for use in this study, I used the Commission's version of the CSRs, split into sub-level CSRs. The Commission divides each CSR into sub-level recommendations, slicing the recommendation into separate and more granular recommendations for different policy areas (Guidi et al., 2023). For example, the 2019 CSR to the Czech Republic to "*Support employment through active labour market policies, the provision of skills, including digital skills, and access to digital learning*" (CSR Czechia, 2019) has been split into a sub-level CSR for active labour market policy (Support employment through active labour market policies,) and one for skills (the provision of skills, including digital skills, and access to digital learning). The RRF reforms by the Member States prescribe a CSR match, but not at the sub-level – the matching process did not encounter any reforms addressing a full CSR; instead, it only matched a sub-level CSR.

To convert the CSRs from 2019 and 2020 for use in this study, identical recommendations for the same Member State are merged. Where the CSRs from 2019 and 2020 for a given Member State overlap – producing fully or near-identical sub-recommendations across the two cycles – the duplicate is counted once, assigning a 'merger ID'. This system avoids an inflated denominator, as I decided to only match each reform once. The feature of repeated recommendations in the European Semester is meant to illustrate that a policy challenge persists, not that it should be addressed each year individually. For instance, Latvia had a CSR recommending investment in sustainable transport in both 2019 and 2020, but implemented only one reform addressing this recommendation. As the reform is only matched once, merging the two identical

CSRs ensures that the score for Latvia reflects that the policy intent of the recommendations was addressed, instead of a score reflecting that Latvia only addressed one of its transport recommendations. Table 1 illustrates examples of sub-level CSRs with policy area labels created by the Commission (European Commission, n.d.) and how I have applied merger identifiers to identical recommendations from 2019 and 2020.

The score of addressed sub-level CSRs reflects the share of attainable sub-level CSRs for each Member State; to ensure that, sub-level CSRs that are not manageable for the purpose if the RRF are excluded. Between two and four sub-recommendations per Member State were excluded, as they could not plausibly be addressed through the reform component of the RRF – for example, the sub-recommendation issued to all Member States in 2020 to "*Front-load mature public investment projects*"(CSR Romania, 2020, p. 8). No implemented reforms addressed this CSR, which is better suited for the investment branch of the RRF or self-standing national policies. Additionally, many Member States had CSRs from 2019 and 2020 recommending budget discipline and management of debt, which most of them did not address: that is expectable, as the performance-based payment system design of the RRF cannot measure what a Member State does not do, i.e., cuts to spending or not issuing further debt.

However, this analysis cannot provide an understanding of the RRF instrument's full engagement with the CSRs, as all the milestones and targets related to investments are not included. This limitation is by design, as the RRF payments to investments are linked to the estimated cost of that investment and essentially reimburse costs once the milestones and targets have been fulfilled. Therefore, the incentive structure and conditionality are less pronounced, as the incentive is the outcome of the investments and not the RRF payment itself. Additionally, this development leads back to a core ambiguity in the governance and implementation design of the RRF: the lack of a legal definition of a reform or investment (Regulation (EU) 2021/241, 2021). It is also not possible to use the CSRs to choose or discard milestones and targets rather than the label 'reform': while their purpose is to recommend structural reforms, several of them encourage public financing. In the end, this research relies on what the Member States, as approved by the Commission and the Council, label as milestones and targets for a reform. The RRF reforms used for this analysis and comparison to the sub-level CSRs thus have a high internal legitimacy, and are the true novelty introduced by the RRF's performance-based payment model.

4.2.2 The implemented reforms under the Recovery and Resilience Facility (RRF)

The available data on implemented reforms is heterogeneous across Member States despite all records being published using the same template, and each Member State's reforms are therefore assessed individually when matched against the relevant sub-level CSRs. Each Member State worked with the Commission to develop its RRF plan, which was subsequently adopted by the Council. The plans detail qualitative or quantitative benchmarks – milestones and targets – for each investment project and reform. When these benchmarks are met for a given reform, the Member State submits the documentation specified in the milestone or target for assessment by the Commission; on positive assessment, the corresponding payment is released, and the fulfilled milestone or target is made publicly available on the dedicated RRF website. This study relies exclusively on milestones and targets related to reforms that have been fulfilled and disbursed before 1 May 2026, four months before the programme deadline at the end of August 2026. Whilst only 65 pct. of milestones and targets related to reforms are fulfilled at the time of writing, Member States have had almost five years to deliver the reforms they planned (Commission, 2021). Reforms, being less resource- and time-intensive than investment projects, were expected to have been front-loaded in the programming period as they were also frequently designed to pave the way for investment projects (ECA, 2025a, p. 10).

The RRF reforms are central to the matching process in the research design; the distribution of milestones and targets can illustrate patterns of planning reforms. The completed milestones, targets, and reforms span a multitude of different policy instruments and approaches; this feature is not considered further in this study. Furthermore, I do not distinguish between reforms with a single milestone or target and reforms with many. Either of these considerations would have introduced complexities to the research design that are ultimately challenged by the primary limitation of the matching process: the inability to assess the depth of reform alignment with the CSRs. Similarly, the research design cannot distinguish between reforms that reflect genuine domestic ownership of the CSR agenda and those that represent surface compliance –

a distinction Jacoby and Hopkin (Jacoby & Hopkin, 2019, p. 24) identify as critical to whether reforms are sustained beyond the conditionality period. Whether the reforms captured in this analysis represent durable structural change, therefore, remains an open question.

Table 1 provides four examples of reforms matched to sub-level CSRs across four different Member States, illustrating both the heterogeneity of the data and the range of information drawn upon in the matching process. The examples show how the same reporting template is used differently across Member States: some include timelines, some concentrate the most substantive description of the reform in the reform description field, whilst others provide the most detail in the milestone or target description. For the CSRs, the matching draws on the text of the sub-level recommendation and the policy area label assigned by the Commission. For the implemented reforms, the matching draws on the reform name, reform description, milestone or target name, and milestone or target description. The column indicating the number of milestones or targets per reform – at least one – reflects the fact that Member States may attach several sequential benchmarks to a single reform, securing payments for partial implementation or intermediate steps rather than only upon full completion. Where several milestones or targets are attached to the same reform, their names and descriptions are combined in the relevant cells, separated by a plus sign (+). Table 1 therefore also illustrates the variation in how Member States have structured the conditionality within their own plans, which is relevant context for interpreting the matching judgements made throughout the first step.

Table 1: Examples of RRF reforms matched to sub-level CSRs

	Finland	Poland	Belgium	Italy
	Merger ID			
CSRs	TRANSPORT	HEALTH	CIRCULAR	JUSTICE
CSRs	**Sub-level CSR**			
CSRs	2019 3.3: and focus investment-related economic policy on sustainable transport, taking into account regional disparities. // 2020 3.5: sustainable and efficient infrastructure	2019 3.6: [Focus investment-related economic policy on] healthcare // 2020 1.2: Improve resilience, accessibility and effectiveness of the health system, including by providing sufficient resources and accelerating the deployment of e-health services.	2020 3.7: the circular economy,	2019 4.2: Improve the effectiveness of the fight against corruption by reforming procedural rules to reduce the length of criminal trials. // 2020 4.1: Improve the efficiency of the judicial system and
CSRs	**Policy area(s)**			
CSRs	Transport;Regional development & local public services // Renewable energy, energy infrastructure & networks;Energy efficiency;Transport	Healthcare	Energy, resources & climate change	Corruption;Justice system
RRF reforms	**Reform name**			
RRF reforms	Roadmap for fossil-free transport	Creating the right conditions for an increase in the number of medical staff	Governance Circular Flanders	Reform of criminal justice
RRF reforms	**Reform description**			
RRF reforms	The objective of this measure is to create a long-term strategic framework with the aim to reduce emissions stemming from the domestic transport sector in Finland. This measure consists of the adoption of a Roadmap for fossil-free transport, presenting a set of measures aiming to achieve emission free transport by 2045.	The objective of the reform is to contribute to improving the matching between needs and availability of medical professionals in Poland. The reform shall consist of initiatives aimed at incentivising young people to take up and pursue medical studies and subsequently practise medicine in Poland. It comprises (i) the introduction of a scheme for granting loans to medical students, including financial incentives to practice in	Circular Flanders is the central platform in Flanders aimed at facilitating the transition to a circular economy in cooperation with industrial partners, knowledge institutions, public administrations, banks and civil society. This reform shall significantly improve and expand the governance of this	The objective of this reform is to shorten the length of criminal proceedings. The reform consists in the adoption of actions that simplify existing procedures and increase courts' productivity.

	Poland after their studies end and (ii) the creation of second-cycle studies for medical emergency practitioners, which shall translate in increased qualifications and revenues for concerned professionals. The reform shall also consist of the introduction of legislation to improve the attractiveness of medical professions and working conditions of medical professionals. This legislation shall increase the flexibility of post-graduate training, including by enabling doctors to be awarded a new certificate of professional competence in various specialised fields. It shall also increase the lowest basic salary for a wide range of medical professionals and reorganise the attribution of certain competences between doctors and specialised medical professionals, emergency medical professionals, nurses and other medical caregivers, after providing appropriate training. The implementation of the reform shall be completed by 30 June 2026.	platform. In that framework, public-private partnerships shall focus on a combination of thematic work agendas, such as circular construction, chemistry/plastics, other product chains in the manufacturing industry, bio-economy and the food chain, and a number of strategic levers (financing, innovation, circular supply, research, awareness). The texts of the roadmap and work agendas are expected by 31 December 2021.	
Nr. of M/T			
2	2 included out of 3	1	1
Milestone/Target Name			
Adoption of the Government resolution LVM/2021/62 to reduce greenhouse gas emissions from domestic transport +"Publication of government decision regarding additional national measures to reduce emissions from domestic transport"	Amendment to the Law on Higher Education and Science and on the Professions of Physician and Dentist to provide a legal basis for financial support from academic year 2021/2022 for students in the field of medicine in Poland +Entry into force of the legislative Act on the profession of paramedic and the self-government of paramedics, which shall introduce the possibility of creating second-cycle	Launch of Steering Group of Circular Flanders	Entry into force of enabling legislation for criminal justice reform

	programmes in the field of preparation for the profession of paramedic		
Milestone/Target description			
The roadmap for fossil free transport shall include instructions and guidelines for the preparation of measures to reduce greenhouse gas emissions from domestic transport by 50% by 2030 compared to 2005 levels. Some of the measures shall be implemented by legislation and others by policy measures.+The measures to be adopted shall follow a comprehensive impact assessment and shall aim at reducing transport emissions by 50% in 2030 compared to 2005. The estimated amount of reductions needed with additional measures is 1.03 megatonnes (as estimated in the baseline scenario April 2020). The estimate shall be updated based on the updated base scenario that shall be presented in autumn 2021.	The law shall enter into force and introduce the possibility of using financial support in the form of a loan for students of paid studies in the field of medicine at a university level. The student shall be able to apply for early repayment of the loan or for an extension of its repayment period. After meeting certain conditions specified in the Act, the student shall be able to apply for a partial or complete remission of the loan for medical studies. Students who take advantage of the support shall be able to apply for a complete remission of the loan upon meeting the following conditions: work after graduation for a period of not less than 10 years within 12 consecutive years counted from the date of graduation, in entities performing medical activities on the territory of the Republic of Poland, which provide health care services financed from public funds, and obtain the title of specialist within the above-mentioned period, in a field of medicine recognised as a priority on the day the physician commences the specialisation training. A person who meets these two conditions shall not have to reimburse the loan for medical studies. Detailed conditions and procedure for	Steering group for the governance of Circular Flanders appointed and Roadmap and Work Agendas for guiding circular economy projects and the development of strategic levers adopted	Enabling legislation which shall include at least the following measures: i) a reviewed notification system, ii) a broader use of simplified procedures, iii) a broader use of electronic filing of documents, iv) simplified rules on evidence, v) the definition of time limits for the duration of preliminary investigation and measures to avoid stagnation in the investigative phase, vi) extension of the possibility to extinguish the crime if damages have been repaid, vii) introduction of a monitoring system at Court level and increase the productivity of criminal courts through incentives to ensure reasonable length

		cancelling the loan shall be specified in the legislative act. +In order to improve the competences of paramedics, a legislative act shall enter into force and enable the creation of second-cycle programmes in the field of paramedic, defined as two-year studies ending with obtaining a master's degree. Obtaining the master's degree shall allow paramedics to gain additional qualifications, which aims to translate into higher ranking in the salary category.		of proceedings and uniform performances across courts.

4.3 The second step - the relationship between the financial incentive and the addressed CSRs

The CSR uptake score derived from the first step serves as the dependent variable in a bivariate analysis, with the size of the financial incentive represented by the RRF envelope as a share of GDP in 2023 as the independent variable. This analysis tests the theorised relationship between financial incentives and reform uptake across the Member States through construct validity, assessing whether the instrument produces the outcomes it was designed to generate. By design, the RRF envelopes are of different sizes across Member States, meaning the incentive structure is asymmetric. As the sample consists of all 27 Member States – constituting 100 pct of the potential population – a scatterplot, Pearson's r, and Spearman's ρ are used to examine the relationship. Both variables are ratio variables, as they are relative measures with equal intervals and a meaningful value of zero: a Member State with a score of 50 pct has addressed twice the share of its sub-level CSRs as one with a score of 25 pct, and a value of zero for either variable carries a substantive meaning (Clark et al., 2021a, p. 324). The scatterplot reveals whether a linear relationship exists between the two variables. Pearson's r indicates the strength and direction of that relationship through a coefficient ranging from -1 to 1, and Spearman's ρ confirms whether the relationship holds when assessed on ranked values rather than raw scores, reducing the influence of any individual outlier (Clark et al., 2021a, p. 333).

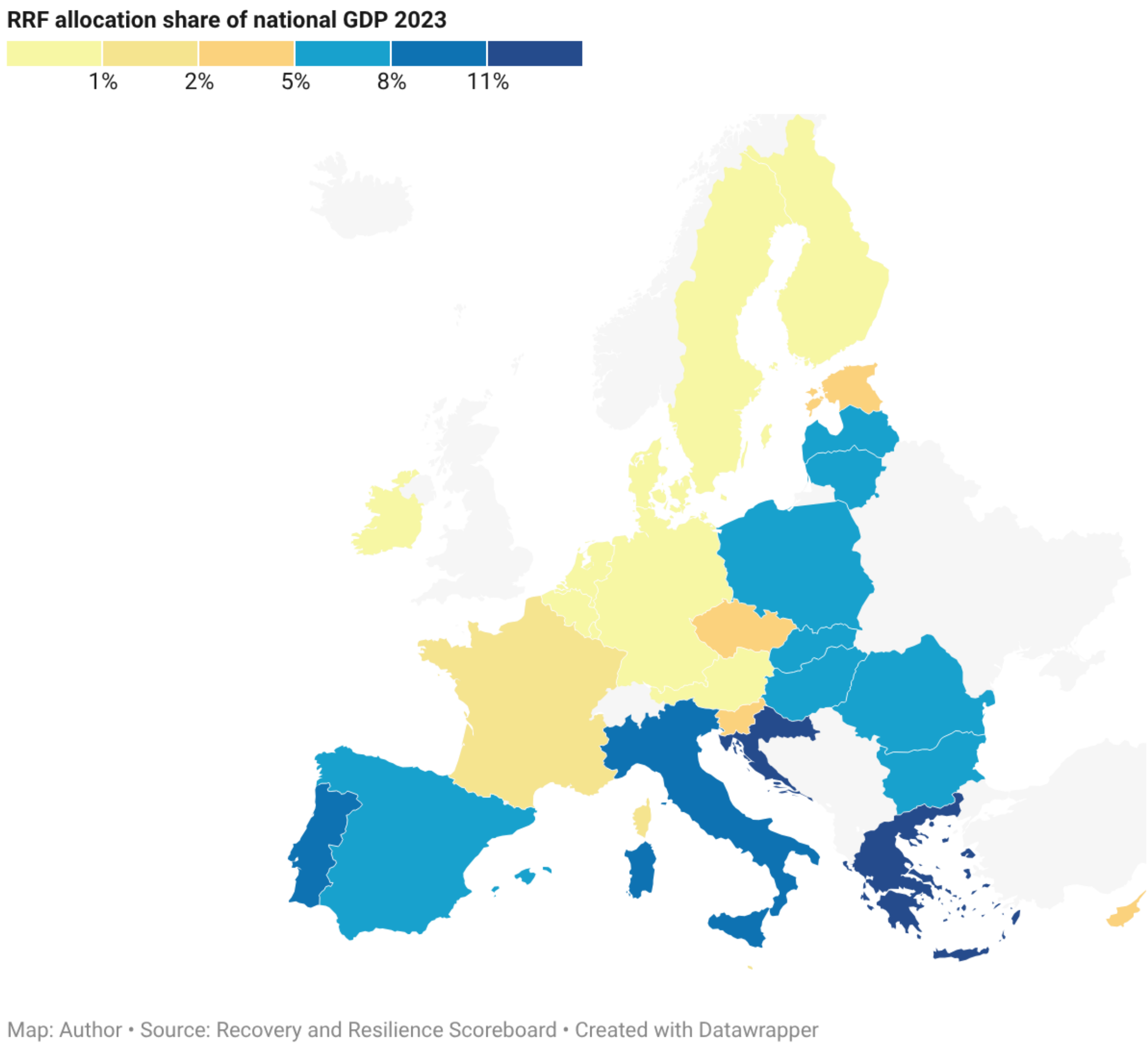
Standardised OLS regressions are applied as a complement to the bivariate analysis to assess the robustness of the observed relationship under different model specifications. The first model fits the CSR uptake score and the RRF envelopes as a share of GDP in 2023. The second model controls for each Member State's total number of CSRs. The score calculation in the first step of this analysis is constructed such that Member States with fewer CSRs face a disproportionately large penalty for missing a single recommendation relative to Member States with many – introducing potential volatility in the dependent variable that is unrelated to the financial incentive. The third model is a robustness check including a measure for economic vulnerability, as the different economic realities of Member States could be a confounding factor to the effect of the financial envelope; however, it is not included in the primary specification, as it partially determines the size of RRF, and its inclusion would therefore introduce an endogeneity problem. A fully specified fourth model includes all three predictors simultaneously, the results of which are interpreted with appropriate caution: with three predictors, the model operates at the boundary of what cross-sectional OLS can reliably support for this sample size (Huntington-Klein, 2025). Other variables such as administrative capacity, political stability or trust in the EU are not considered, as they are less relevant to the theoretical framework. The following sections elaborate on the variables underlying the second step of the research design: the financial incentive measured by the size of the RRF envelope and the variables for control.

4.3.1 The national RFF envelope

The incentive structure in the RRF is inherently asymmetric, meaning that the principal-agent dynamics theorised above apply to varying degrees across the 27 Member States. The RRF determined a ceiling for grants and loans for each Member State using an allocation key that takes relative poverty and unemployment rates across regions into account. The result was considerably larger allocations as a share of GDP for Eastern and Southern European Member States than for their Northern and Western counterparts (Barbero et al., 2022, p. 2). The asymmetry is further compounded by the option for a voluntary loan component. As shown in Table 2, only 12 Member States opted to take up loans before the deadline in August 2023, whilst uptake on the grant component was universal. Two measures of RRF funding relative to GDP were initially considered – as a share of GDP in 2023 and as a share of cumulative GDP across 2021–2026. As the latter relies on a forecast for 2026, it introduces measurement uncertainty absent from the

2023 figure, which is published directly by the Commission on the official website. Given that both measures produced substantively identical results across all specifications, GDP 2023 was retained as the preferred measure. Figure 1 illustrates the resulting distribution of RRF funding across Member States as a share of GDP in 2023. The Member States coloured yellow receive less than 5 pct of their GDP in 2023, while the blue colour marks the 12 Member States receiving the most. The extreme cases are Ireland, Luxembourg, and Denmark, receiving less than 0.5 pct and Italy, Croatia, and Greece receiving more than 9 pct.

Figure 1: The RRF envelopes range from 0,2 pct to 16 pct of GDP in 2023



However, it is important to keep in mind that the theorised relationship between reform uptake and the RRF envelope is indirect, as this variable does not illustrate the amount of funding available for reforms, but investments and reforms combined. This connects to the quality of data – while using archival public data is reliable and of high quality, key variables for this research design are missing, such as the payout values per milestone or target. Had this been present, the analysis could have presented stronger conclusions on the effect of funding, especially across the policy areas, by bivariate or multivariate analysis. Alternatively, this data would have made it possible to give special attention to the reforms receiving the most funding and which Member States prioritised the compensation for reforms rather than reimbursement of investments. Neither conclusion can be made based on this thesis, as there is no ring-fencing of funding for reforms or a set payment value for reforms.

4.3.2 Economic vulnerability and reform pressure

As the dependent variable is constructed as the share of sub-level CSRs addressed through reforms, the number of unique CSRs a Member State receives mechanically enters the denominator, making it necessary to control for this to avoid conflating variation in CSR uptake with variation in the size of each Member State's CSRs. This is particularly consequential for Member States with few sub-level CSRs, where failing to address a single recommendation produces a disproportionately large drop in the percentage score compared to Member States with more CSRs – introducing potential volatility in the dependent variable that is unrelated to the financial incentive. For example, Ireland has 20 unique sub-level CSR recommendations, while Bulgaria has 31: not addressing one recommendation decreases Ireland's score by 5 pct, while the same decreases Bulgaria's score by only 3 pct. Additionally, a higher number of unique sub-level CSRs indicates that either the pandemic has revealed challenges prone to the relevant Member State, or indicates a previous unwillingness to implement structural reforms compared to Member States with fewer recommendations. This last consideration is connected to whether the Member State is relatively more economically vulnerable, which either drives more recommendations or is a result of an unwillingness to reform, as Alesina and Drazen (1991) theorise.

The last variable measures economic vulnerability using GDP per capita in 2023 – a deliberately simplified measure that mirrors the year used for the RRF envelope calculation, avoiding the need to translate figures across price levels. The choice of 2023 is deliberate: by then, the immediate pandemic distortions had subsided, and domestic stimulus packages had stabilised economies, whilst RRF payments were still limited enough not to have substantially altered the underlying economic picture across Member States. A more comprehensive index –combining debt levels, deficits, unemployment, and private savings as in Armingeon and de la Porte (2022, p. 146) – would more precisely capture economic vulnerability, but would introduce data from different years and a complexity that is not warranted given the variable's role as a robustness check rather than a primary predictor. This is the case as structural economic conditions drove the allocation of RRF envelopes – primarily pre-existing unemployment and GDP per capita rather than pandemic severity itself (Armingeon et al., 2022; de la Porte & Jensen, 2021). Given this endogeneity to the independent variable, GDP per capita in 2023 is included not as a control but as a robustness check to assess whether structural economic differences between Member States independently explain variation in CSR uptake beyond the size of the financial incentive.

4.4 The complete dataset

Table 2 provides an overview of all the data material used for the second step, along with results from the first step. While the data on grants and loans, along with the cumulative payments before 1 May 2026, are not part of the second step, they provide some context for this study of the RRF. The last column, 'score', is the dependent variable, describing the uptake of sub-level CSRs through reforms for each Member State.

Table 2: Overview of the CSRs, RRF reforms, RRF envelope and GDP 2023

Country	RRF envelope (bn €)	Share of GDP 2023 (%)	GDP per capita 2023 (€k)	Grant planned (bn €)	Loan planned (bn €)	Disbursement 1 May (bn €)	Reforms	M&T	Unique CSRs	Eligible CSRs	Addressed CSRs	Not adressed CSR	Score 1)/2) (%)
Austria (AT)	3,96	0,84%	52,3	3,96	-	3,33	28	56	21	18	15	3	83%
Belgium (BE)	5,26	0,88%	51,1	5,03	0,23	3,07	29	37	23	19	12	7	63%
Bulgaria (BG)	6,17	6,52%	14,7	6,17	-	3,27	33	67	31	29	16	13	55%
Croatia (HR)	10,04	12,86%	20,5	5,79	4,25	4,91	60	113	26	24	21	3	88%
Cyprus (CY)	1,02	3,26%	33,9	1,02	-	0,59	40	49	29	27	22	5	81%
Czech Republic (CZ)	8,7	2,74%	29,3	8,36	0,34	6,77	44	74	18	16	11	5	69%
Denmark (DK)	1,63	0,43%	62,9	1,63	-	1,27	4	6	9	7	3	4	43%
Estonia (EE)	0,95	2,50%	28,1	0,95	-	0,63	13	25	16	13	6	7	46%
Finland (FI)	1,95	0,71%	48,9	1,95	-	1,12	12	21	15	12	7	5	58%
France (FR)	40,27	1,43%	41,3	40,27	-	34,13	16	29	20	17	10	7	59%
Germany (DE)	30,32	0,72%	50,7	30,32	-	19,76	17	28	17	15	7	8	47%
Greece (EL)	35,95	15,96%	21,3	18,22	17,73	23,44	63	112	19	17	14	3	82%
Hungary (HU)	10,43	5,30%	20,6	6,51	3,92	0,92	0	0	23	20	0	20	0%
Ireland (IE)	1,15	0,23%	99,1	1,15	-	0,68	9	32	20	17	8	9	47%
Italy (IT)	194,38	9,12%	36,3	71,78	122,6	153,17	54	125	29	27	16	11	59%
Latvia (LV)	1,97	5,00%	21	1,97	-	1,09	20	38	18	15	9	6	60%
Lithunia (LT)	3,84	5,20%	25,9	2,29	1,55	2,7	23	74	18	16	10	6	63%
Malta (MT)	0,33	1,60%	37,8	0,33	-	0,22	28	56	18	16	11	5	69%
Spain (ES)	102,56	6,84%	31	79,85	22,71	71,36	82	133	28	26	21	5	81%
Netherlands (NL)	5,44	0,51%	58,7	5,44	-	3,07	20	32	16	14	6	8	43%
Slovenia (SI)	2,14	3,34%	30,2	1,61	0,53	1,54	22	46	22	19	11	8	58%
Poland (PL)	54,72	7,31%	20	25,28	29,44	26,93	32	67	21	18	13	5	72%
Romania (RO)	21,4	6,60%	16,9	13,57	7,83	10,72	39	79	24	20	14	6	70%
Portugal (PT)	21,91	8,18%	25,6	16,32	5,58	14,92	33	66	20	17	11	6	65%
Luxembourg (LU)	0,24	0,30%	123	0,24	-	0,09	7	15	18	16	7	9	44%
Sweden (SE)	3,45	0,64%	50,5	3,45	-	1,65	10	10	12	10	5	5	50%
Slovakia (SK)	6,41	5,21%	22,6	6,41	-	3,96	52	69	23	21	17	4	81%
Total	576,6	-	-	359,88	216,71	395,3	790	1459	-	-	-	-	-

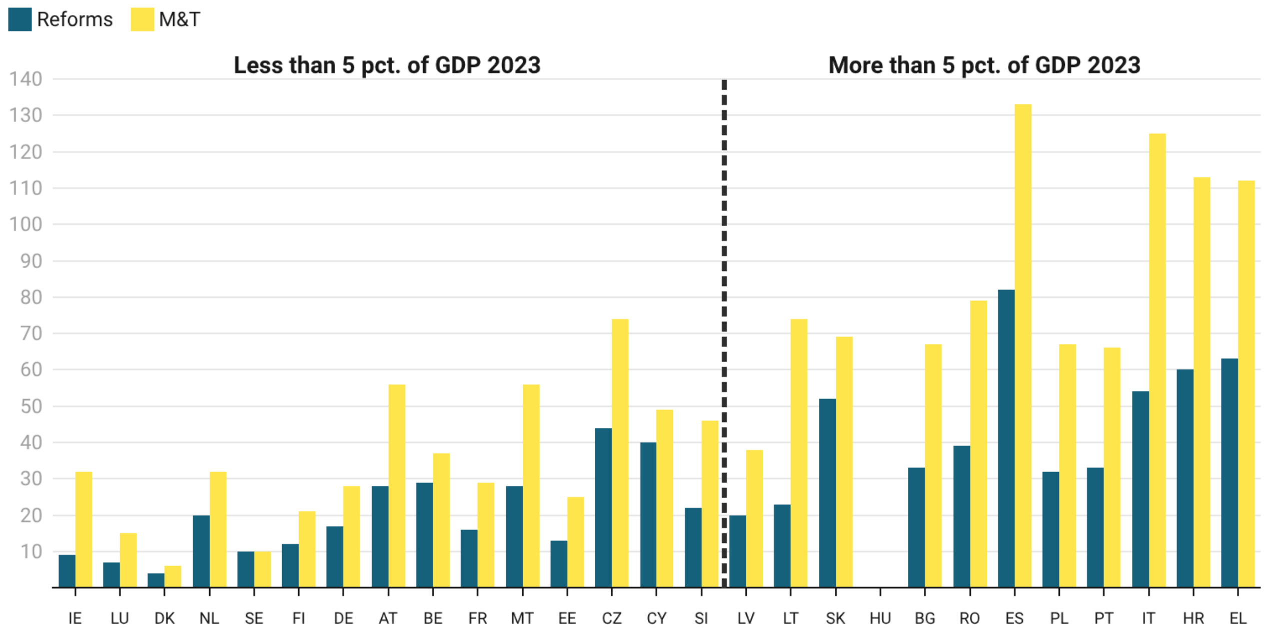
5. Cash leads to some reforms: But not all of them

5.1 The reforms under the RFF are very country-specific

The first step of the research design reveals considerable variation in the number of implemented reforms across Member States, with the distribution in Figure 2 indicating that Member States with envelopes larger than 5 pct of their GDP in 2023 have broadly opted for more reforms than those with smaller envelopes. The heterogeneity extends beyond volume: both the reforms themselves and their associated milestones and targets differ vastly across Member States, reflecting the ambiguous requirements of the RRF regulation combined with the novelty of the instrument – it is the first time that Member States' administrations have planned and implemented reforms through an EU programme (ECA, 2023, p. 28). This ambiguity is itself a manifestation of the monitoring challenge identified in the principal-agent framework: Member States negotiated separately with the Commission at the planning stage, which produces meaningful variation in how Member States have structured their conditionality and, by extension, in how they have managed the bonding costs of the instrument (ECA, 2025a, p. 11).

As Figure 2 illustrates, Sweden is the only Member State to attach a single milestone or target to each reform, whilst Italy – with the largest RRF envelope – averages 2.3 milestones and targets per reform, indicating a preference for sequential conditionality that secures payments for partial or intermediate implementation steps. The variation also extends to how the payment model is applied in practice: an example is Lithuania, where individual milestones and targets can constitute sub-reforms within a broader reform package, i.e., 11 milestones and targets on reforms to different parts of the educational sector, while the overall reform is to modernise general education. Meanwhile, Ireland attaches six milestones and targets to a single reform consisting of annual incremental increases over 6 years to a carbon tax.

Figure 2: The number of reforms addressing CSRs varies across the 27 Member States



5.1.1 Signs of the moral hazard problem?

Across all 27 Member States, 122 fulfilled reforms were discarded as not fitting the CSRs – 93 by the discretion of the reporting Member State, and 29 based on the coding criteria developed throughout the matching process. The former category consists of reforms that Member States themselves did not reference to any CSR in their reporting, most of which are reforms related to the implementation of the RRF itself or other EU policy documents. The latter consists of reforms where a CSR reference was provided, but where no reasonable match could be identified upon assessment against all sub-level CSRs for that Member State. This broader assessment against all sub-level CSRs – rather than only the referenced ones – is necessary because Member States have an incentive to report reforms against as many CSRs as possible, or against a different CSR than the most accurate fit, since the RRF regulation requires that plans address all or a significant subset of the recommendations. This behaviour is consistent with the moral hazard problem identified in the principal-agent framework: because the reward is tied to documentation rather than to an independent assessment of reform quality, the agent has a rational incentive to present measures as CSR-responsive even where the substantive alignment is weak.

Table 3 illustrates three examples of discarded reforms from Cyprus, Spain, and Lithuania that reference a CSR. In the case from Cyprus, the reform “Online platform for trade and information symmetry in the fresh produce supply chain" is a very narrow initiative for a very specific sector and fails to describe how it adds to the objective of any CSR, let alone the reported CSR, despite its broad coverage across the green transition, environment, liquidity to businesses, digitalisation, and transport. For the case from Spain, a "Law for Sports", it was simply not possible to find a suitable match among Spain's CSR. The reform only describes the objective, which does not match Spain's CSRs, and does not explain any strategy to chase that objective. Finally, the Lithuanian reform "An Efficient Public Sector" only addressed the skills of civil servants, instead of seeking to contribute to active labour market policy as recommended. The objective of the reforms is to improve public administration – a field in which Lithuania has no CSR in.

Table 3: Examples of reforms that did not fit any CSRs

Country	Cyprus	Spain	Lithuania
Reform name	Online platform for trade and information symmetry in the fresh produce supply chain	Law for Sports	An Efficient Public Sector
Reform description	The objective of the reform is to address long-standing drawbacks of the fresh produce supply chain. The reform consists in the entry into force of a law on unfair practices in transactions in the local fresh produce market. A Platform shall be offered to Producers' Organisations in the fresh produce market.	The objective of this legislative measure is to ensure health and safety in the practice of sports at all levels, to include gender equality, accessibility and social cohesion aspects in the regulation of sports, to modernise sports organisations and infrastructures through digitisation and through their ecological transformation and to promote the internationalisation of the sector. The legislation shall adapt the organisational structures of sport to the current challenges that it faces, taking account of lessons learned from the pandemic. The implementation of the measure shall be completed by 31 December 2022.	This reform aims to reform the civil service by improving administrative processes, human resources management and strengthening customer-orientation in the public sector. Implementation of this reform shall require the adoption of the relevant legislation, including amendments to the Law on the Civil Service. This reform consists of two sub-measures: (1) modernization of human resources management system in the public sector (sub-measure 1); (2) establishment of a centralised training system for developing competences in the public sector (sub-measure 2).
Milestone / target name	Law on unfair practices in transactions in the local fresh produce market	Entry into force of the law for Sports	Strategic guidelines and training modules
Milestone/ target description	Entry into force of a new law which shall address unfair trading practices in transactions in the local fresh produce market, such as unilateral and	The Law shall foster health and safety in the practice of sports at all levels, gender equality, social inclusion and accessibility, the promotion of the international	Strategic guidelines for the long-term training and competence development of public sector employees and implementation plan of the strategic guidelines shall be adopted

	retroactive amendments to contracts, last minute cancellations, payment deadlines longer than 30 days, payment for damaged or unsold products and other actions affecting those involved in the production and distribution chain of agricultural products.	dimension of the model and the modernization of organisations, and infrastructures through respect for the environment and digitization.	by the Government of Lithuania. Training modules for strengthening the competencies of public sector employees shall be developed by the Public Management Agency. The following training modules shall be developed: 1) digital competencies; 2) financial-analytical competencies; 3) leadership competencies.
CSR reference	2019.CSR4; 2020.CSR3: Secure adequate access to finance and liquidity, especially for small and medium-sized enterprises. Front-load mature public investment projects and promote private investment to foster the economic recovery. Focus investment on the green and digital transition, in particular on clean and efficient production and use of energy, waste and water management, sustainable transport, digitalisation, research and innovation.	2019.CSR3; 2020.CSR3: Focus investment-related economic policy on fostering innovation, resource efficiency, and energy efficiency, upgrading rail freight infrastructure and extending electricity interconnections with the rest of the Union, taking into account regional disparities. Enhance the effectiveness of policies supporting research and innovation.	2020.CSR2: Mitigate the impact of the crisis on employment. Increase the funding and coverage of active labour market policy measures and promote skills.

5.1.2 The case of Hungary

An outlier among the 27 Member States is Hungary, where zero milestones or targets have been deemed fulfilled by the Commission, and which is therefore excluded from the correlation analysis. Hungary's absence from the results is not primarily a failure to implement reforms in line with the CSRs, but rather a consequence of unmet rule of law conditions embedded in the RRF regulation. Thinus (2025) finds that rule of law protection in the RRF is treated primarily as a means of ensuring sound financial management and a credible investment climate rather than as an end in itself. Credible protection of EU funds against corruption, fraud, and conflicts of interest is required as a precondition for any payment. Milestones related to these protections have been designated as 'super-milestones', meaning no other milestone or target can be verified as fulfilled until sufficient control and audit systems have been approved. Hungary's investment plan, subject to intense negotiations from May 2021 to December 2022, ultimately contained almost 80 milestones and targets on rule of law protection, among them 27 ‘super-milestones’ (Thinus, 2025, p. 888). The previous government under Viktor Orbán refused to implement these, thus blocking the full envelope. The new government under Peter Magyar entered negotiations with the Commission in May 2026 to adjust these milestones to unblock the funding before the deadline in August 2026 (Zsiros, 2026).

In principal-agent terms, Hungary represents a residual loss: the principal's incentive through conditionality failed entirely to produce the intended alignment. While Hungary illustrates the limits of the conditionality structure, the data do not reflect unwillingness to implement the planned reforms themselves or address the corresponding CSRs across the full spectrum of policy areas, but instead the reluctance to see through the ‘super milestones’. For that reason, Hungary is excluded as a structural outlier rather than as a poorly performing case – the other Member States with a similarly sized envelope record a CSR uptake score of at least 63 pct.

5.2 Larger envelopes predict a higher score of addressed CSRs

5.2.1 Scatterplot and bivariate analysis

The analysis reveals a positive relationship between the size of RRF funding relative to GDP and the share of CSRs addressed through reforms, providing empirical support for the principal-agent framework's central proposition: that a credible financial incentive produces measurable alignment between the agent's behaviour and the principal's goals. As illustrated in Figure 3, Member States receiving larger envelopes as a share of GDP 2023 tend to display higher rates of CSR uptake, with the fitted regression line indicating a consistent upward trend across the distribution. Whilst there is notable variation around the trend line – with Member States such as Austria and Cyprus performing above expectation relative to their funding size, and Bulgaria and Estonia performing below – the overall pattern is sufficiently clear. The grey shaded area marks the 95 pct confidence interval around the fitted regression line, illustrating the uncertainty of the best estimate of the relationship. The area is broader when there are fewer observations, which is especially evident when the independent variable is above 10 pct, meaning the fitted line is more uncertain.

Figure 3: Scatterplot of score of addressed CSRs and RRF funding as share of GDP 2023

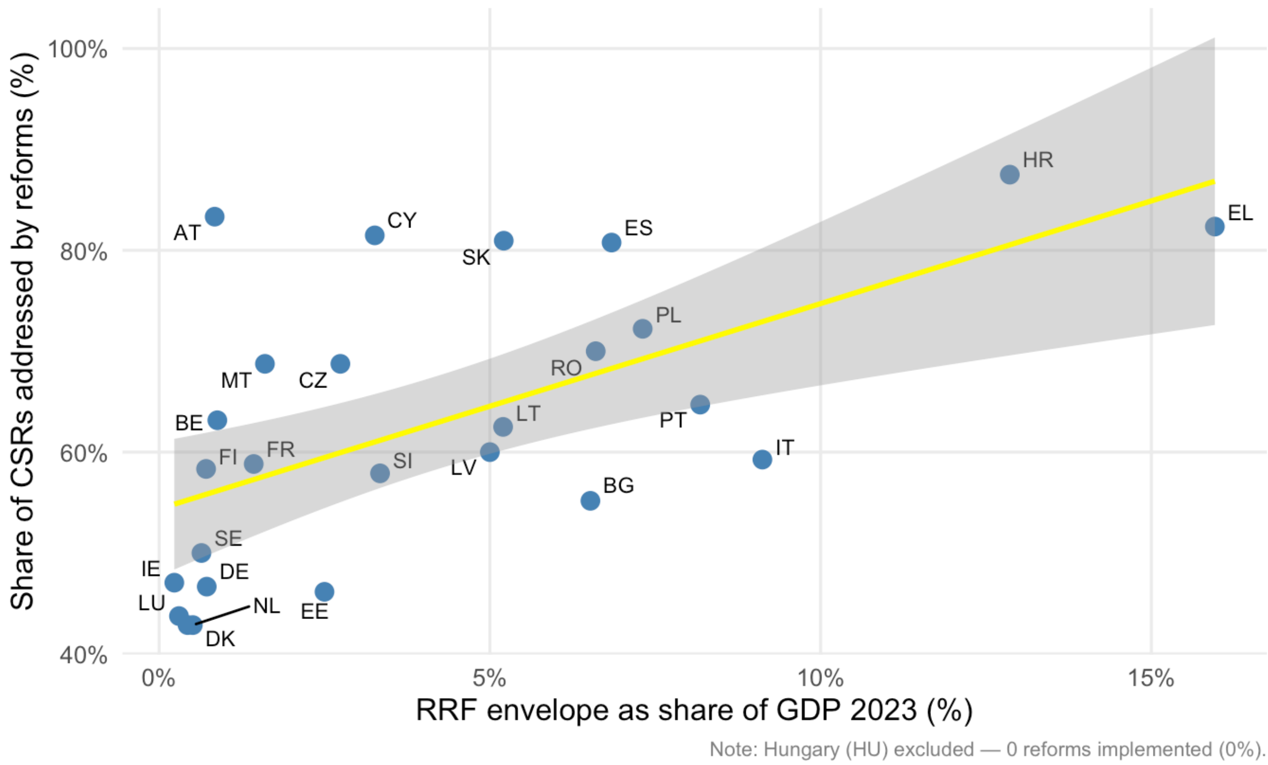


Table 4 reports the results of both Pearson's product-moment correlation (r) and Spearman's rank correlation (ρ), which corroborate the visual pattern. Pearson's r = 0.60 (p = 0.001, 95% CI [0.28, 0.80]) indicates a moderate positive linear relationship, whilst Spearman's ρ = 0.66 (p < 0.001) confirms that the relationship holds when assessed based on ranked values rather than raw scores, reducing the influence of any individual outlier. When Hungary is included in the analysis, both coefficients weaken noticeably – Pearson's r drops to 0.41 (p = 0.032) and Spearman's ρ to 0.57 (p = 0.002) – confirming that Hungary acts as an influential outlier that suppresses the overall relationship. The consistency across both methods and sample compositions strengthens confidence that the observed association is not an artefact of the distributional properties of the data and provides a foundation for further analysis.

Table 4: Bivariate correlations between RRF envelope as a share of GDP 2023 and score of addressed CSRs

	Pearson's r	**Spearman's ρ**
Coefficient	0.60	0.66
p-value	0.001	<0.001

95% CI	[0.28, 0.80]	—
N	26	26

Note: Hungary (HU) excluded — 0 reforms implemented (0%). Spearman's p-value based on approximate method due to ties in the data. No confidence interval reported for Spearman's ρ.

5.2.2 Controls and robustness checks

The positive relationship between RRF funding and CSR uptake identified in the bivariate analysis holds when examined through an OLS regression, providing further support for the argument that financial incentives encourage Member States to address CSRs. Model 1 of Table 5 estimates the relationship between RRF funding as a share of GDP 2023 and CSR uptake without controls, and finds a positive and statistically significant effect ($\beta = 0.599$, $p < 0.01$), indicating that RRF funding alone explains approximately 36 pct of the variation in CSR uptake across Member States ($R^2 = 0.359$, adjusted $R^2 = 0.332$). Whilst this baseline model does not account for confounding factors, it establishes that the financial incentive variable has a substantively meaningful and statistically robust association with the dependent variable in its simplest specification, consistent with the correlation results reported in Table 4.

Table 5: Standardised OLS regression: score of addressed CSRs and RRF envelope

	Dependent variable: Score of addressed CSRs			
Model:	**(1)**	**(2)**	**(3)**	**(4)**
RRF envelope / GDP 2023	0.599***	0.444**	0.419**	0.299
	(0.163)	(0.177)	(0.200)	(0.204)
Number of unique CSRs		0.328*		0.302*
		(0.177)		(0.175)
GDP per Capita 2023			−0.298	−0.261
			(0.200)	(0.193)
Constant	0.000	0.000	0.000	0.000
	(0.160)	(0.153)	(0.156)	(0.150)
Observations	26	26	26	26

R^2	0.359	0.443	0.416	0.486
Adjusted R^2	0.332	0.395	0.365	0.416

*Note: Standardised coefficients (β) reported. Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01. Hungary (HU) excluded — 0 reforms implemented (0%). Model (1): RRF envelope only. Model (2): RRF envelope + number of unique CSRs. Model (3): RRF envelope + GDP per capita 2023. Model (4): RRF envelope + number of unique CSRs + GDP per capita 2023. Models (3) and (4) should be interpreted as robustness checks.*

The financial incentive effect persists when controlling for properties of the dependent variable, with Model 2 of Table 5 introducing the number of unique CSRs. The model serves as a control to account for the fact that Member States with fewer sub-level CSRs face greater score volatility for each unaddressed recommendation. The financial incentive variable nonetheless retains a positive and statistically significant standardised coefficient ($\beta = 0.444$, $p < 0.05$) and remains the stronger predictor within this specification compared to CSR volume ($\beta = 0.328$, $p < 0.1$). Together, the two predictors explain approximately 44 pct of the variation in CSR uptake across Member States ($R^2 = 0.443$, adjusted $R^2 = 0.395$), suggesting that whilst the financial incentive is the dominant explanatory factor, reform pressure independently contributes to uptake — consistent with the principal's goal of addressing structural challenges across Member States.

To assess whether the financial incentive effect reflects economic vulnerability rather than incentive response, Model 3 of Table 5 replaces the CSR volume control with GDP per capita in 2023. As noted in the methods section, GDP per capita partially determines the size of RRF allocations through the allocation formula, meaning it is not a fully independent control variable – its inclusion should therefore be interpreted as a robustness check. The financial incentive variable remains the dominant predictor and retains a positive and statistically significant standardised coefficient ($\beta = 0.419$, $p < 0.05$), whilst GDP per capita is negative and statistically insignificant ($\beta = -0.298$, $p > 0.1$), indicating that baseline economic wealth does not independently explain variation in CSR uptake once the size of the financial incentive is accounted for. This is consistent with the principal-agent framework's expectation that it is the material incentive – rather than the structural economic condition of the agent – that drives alignment with the principal's goals: wealthier Member States do not systematically address a higher share of their CSRs through the RRF, whilst poorer ones do not systematically underperform once the size of their financial allocation is held constant. The consistency of the financial incentive coefficient across Models 1 to 3 – remaining significant and positive with comparable magnitude – strengthens confidence that the observed relationship between the RRF envelope and the CSR uptake is indicative of a genuine incentive effect rather than a reflection of structural economic differences between Member States ($R^2 = 0.416$, adjusted $R^2 = 0.365$).

Whilst the preceding models consistently support the argument that financial incentives encourage Member States to address CSRs, the fully specified model introducing all three predictors simultaneously produces a more ambiguous result that warrants discussion. Model 4 of Table 5 includes the RRF envelope as a share of GDP 2023, number of unique CSRs, and GDP per capita 2023 as predictors, and finds that the financial incentive variable loses statistical significance ($\beta = 0.299$, $p > 0.1$), whilst the number of unique CSRs retains a marginal positive effect ($\beta = 0.302$, $p < 0.1$). Two factors limit the conclusions that can be drawn from this result. First, as discussed above, GDP per capita is partially constitutive of the RRF allocation formula, meaning its inclusion as a control likely absorbs variation that legitimately belongs to the financial incentive variable, artificially deflating its coefficient. Second, with only 26 observations and three predictors, the model potentially operates at the boundary of what cross-sectional OLS can reliably support, meaning that the loss of significance more plausibly reflects insufficient statistical power than a genuine absence of effect (Huntington-Klein, 2025). Crucially, the standardised coefficient on the financial incentive variable remains positive and the largest among the three predictors across all four specifications in Table 5, and the model explains approximately 49 pct of variation in CSR uptake ($R^2 = 0.486$, adjusted $R^2 = 0.416$) – suggesting that whilst the financial incentive contributes meaningfully to explaining variation in agent behaviour across Member States, a substantial share of that variation remains unaccounted for, consistent with Jensen and Meckling's (1976, p. 308) observation that some divergence between principal and agent will always persist.

5.3 Energy, education, and transport are favoured policy areas

Whilst the preceding analysis has established that no Member State addressed all its sub-level CSRs through implemented reforms, it does not illuminate whether the residual gap between the principal's preferences and the agents' behaviour is concentrated in particular policy areas of the CSRs. Examining the distribution of reforms across policy areas offers a more granular picture of where the RRF's incentive and conditionality model has been most and least effective. Table 6 provides an overview of reforms matched to unique sub-level CSRs across seventeen policy area categories, compiled based on the policy areas assigned by the Commission; sub-level CSRs can have multiple areas assigned, meaning some are counted more than once. Only five categories are present across all 27 Member States: business environment, energy, education and skills, healthcare, and transport. The most striking finding is the contrast between the highest-performing categories – transport at 83 pct of sub-level CSRs addressed and energy at 82 pct – and the lowest-performing, with business environment at 37 pct, and finance at 40 pct. The following sections examine what this uneven distribution reveals about the nature of the residual loss across the principal–agent relationship in the RRF.

The highest rates of addressed CSRs are found in energy and education, with energy accounting for the largest absolute number of reforms – 147 across all 27 Member States at a rate of 82 pct, with Hungary, Latvia, Malta, and Luxembourg as exceptions. The strong performance of the energy category is partly attributable to the adoption of the RepowerEU framework in 2022, through which 300 billion euros were ring-fenced for RepowerEU plans delivered via the RRF in response to Russia's invasion of Ukraine and the resulting energy crisis (Wójtowicz, 2024, p. 161). Reforms implemented under RepowerEU plans are included in this analysis, and the high uptake in this category therefore reflects a combination of both the RRF incentive model and a stronger alignment between the principal and the agents’ interests due to the energy crisis. Transport, with 83 pct of sub-level CSRs addressed across 27 Member States despite a comparatively modest 73 reforms in absolute terms, and education, offer a more ’usual’ application of the conditionality structure. Education and skills yielded 102 reforms across 23 Member States at a rate of 72 pct, with the Netherlands, Hungary, France, and Denmark as exceptions. By contrast, several policy categories show markedly lower rates of addressed CSRs, and the following sections examine what drives this divergence.

The lowest rates of addressed CSRs are found in the areas of business environment and finance, where the incentive and conditionality structure of the RRF appears to have been least effective. A special report by the European Court of Auditors (2025b) finds that business environment CSRs were largely addressed outside the RRF framework altogether, implying that Member

States accommodated these recommendations through domestic channels to some degree rather than in response to the conditional financial incentive. There is a significant overlap in the sub-level CSR coding between business environment and finance, meaning the Court of Auditors' finding applies more broadly across these two categories. While an interest alignment and domestic handling are part of explaining the low uptake across these categories, some missing matches are an indication of a residual loss.

A final finding tentatively supports the conclusions of Deroose and Griesse (2014, p. 2) from studying the implementation of the CSRs: policy areas not backed by EU-rules with enforcement powers are less amenable to the reform-side of the financial incentive of the RRF. As the whole system of RRF reforms and the CSRs is about voluntary national reforms, the question of EU competence is not central – but housing, pension, and taxation are especially rooted in national competence, and thus further into the agent's own arena. In addition to the competence question, pensions and taxation are characterised by being macroeconomically important for national budgets and very salient politically (Deroose & Griesse, 2014; Efstathiou & Wolff, 2018). Sub-level CSRs on housing, pension, and taxation cover only 6, 16, and 15 Member States, respectively, meaning nothing can be concluded in general. The rates of addressed CSRs for these areas are relatively low, at 57, 53, and 55 pct, respectively. The same seems to be true for recommendations on state-owned enterprises in the category for Single Market: only four out of 7 Member States addressed the relevant sub-level CSRs, while Cyprus, Portugal, and Slovenia did not. In this light, the incentives were simply not adequate for swaying national elites in this direction, and the unaddressed CSRs in these areas represent a residual loss to the principal that the conditionality model proved insufficient to correct.

Another explanation other than national competence and political salience could be the complicated nature of planning and executing, especially housing, pension, and tax reforms, which may have been beyond the deadlines and timeframe of the RRF (ECA, 2025a, p. 39). This is a more legitimate version of the moral hazard problem: these types of reforms were not fit for the RRF planning, not due to unwillingness from Member States, but due to the standards of legislative procedures. Regardless, some reforms were done in these areas: for example, Spain completed four milestones for a pension reform called "Maintenance of the purchasing power of pensions, alignment of the effective retirement age with the statutory retirement age, adaptation of the calculation period for the calculation of the retirement pension to new careers and replacement of the sustainability factor by an intergenerational equity mechanism", which can be interpreted as an honest attempt at addressing the CSR to "(…) *preserve the sustainability of the pension system*"(CSR Spain, 2019). In the area of housing policy, the Netherlands did a reform called "Phasing out the tax exemption for gifts to finance home purchases" to address the CSR "*Reduce the debt bias for households and the distortions in the housing market, including by supporting the development of the private rental sector*" (CSR Netherlands, 2019). For the taxation area, recommendations to shift tax away from labour or broaden the tax base were generally not pursued; instead, reforms were done connected to improving tax compliance or combatting aggressive tax planning.

Excluding Hungary makes the largest relative difference in the policy category of justice system, increasing the score of addressed sub-level CSRs from 67 to 80 pct. The category contains sub-level CSRs in the policy areas of the justice system and corruption. Hungary's three sub-level recommendations in this area are on reinforcing the anti-corruption measures, strengthening judicial independence and ensuring that any emergency measures are proportionate and in line with European and international standards (CSR Hungary, 2020; 2019). However, the results in Table 6 do not illustrate that Hungary was unwilling to address these CSRs despite the financial incentive; as explained in a dedicated section, Hungary implemented zero RRF reforms due to the political opposition to completing the 'super-milestones' relating to the required audit and control systems to receive payments from the RRF. Table 6 shows the results from excluding Hungary in brackets next to the result including all 27 Member States.

Table 6: Overview of reforms on policy areas in 17 categories

Category	Policy area(s)	MS (excl. Hungary)	Unique CSRs	Addressed CSRs (%)	RRF Reforms
Business environment	Business environment (incl. SME policies)	27 (26)	71 (69)	37% (38%)	57
Energy	Energy, resources & climate change Renewable energy, energy infrastructure & networks; Energy efficiency;	27 (26)	38 (36)	82% (86%)	147

Education and skills	Education; Skills & life-long learning; Vocational Education and Training & adult learning; Non-discrimination and equal opportunities; Early childhood education & care	27 (26)	50 (47)	72% (77%)	102
Healthcare	Health & long-term care; Healthcare	27 (26)	39 (38)	79% (82%)	63
Transport	Transport	27 (26)	30 (29)	83% (86%)	73
R&D&I	Research & innovation	27 (26)	33 (32)	70% (72%)	41
Finance	Access to finance; Financial services; Financial services and financial stability; Access to finance & growth financing (incl. CMU); Private sector debt & insolvency framework; Anti-money laundering	26 (25)	58 (57)	40% (40%)	27
Single Market	Competition in services; State-owned enterprises; Single market, competition & state aid	15 (14)	20 (19)	55% (58%)	11
Labour market and wages	Active labour market policies, incentives to work & labour market participation; Wages & wage setting; Functioning of the labour market, including framework for labour contracts & undeclared work;	25 (24)	55 (52)	53% (56%)	49
Digitalization	Telecom, postal services & local public services; Digitalisation of public administration and public services; Digital connectivity, infrastructure & market functioning; Digitalisation of businesses	23 (22)	35 (34)	69% (71%)	80
Poverty	Poverty reduction & social inclusion;	21 (20)	36 (34)	64% (68%)	49
Pension	Long-term sustainability of public finances, inc. Pensions; Pension systems and active ageing	16	17	53%	34
Environment	Environmental policy & resource management	16 (15)	24 (22)	75% (82%)	54
Taxation	Reduce the tax burden on labour; Tax administration, tax evasion & tax avoidance	15 (14)	20 (19)	55% (58%)	22
Public administration	Public administration; Digitalisation of public administration and public services; Public procurement & concessions	14 (14)	28 (27)	71% (74%)	71

Budget and fiscal frame-work	Fiscal policy & fiscal governance; Budgetary framework and fiscal governance	13	17	82%	41
Justice sys-tem	Justice system; Corruption	10 (9)	18 (15)	67% (80%)	31
Housing	Housing	6	7	57%	11

6. Cash for reform: what worked, what did not, and what comes next?

The analysis finds a statistically significant relationship between the relative size of the RRF envelope and the share of sub-level CSRs addressed through reforms in each Member State. The theoretical framework poses this not merely as a correlation, but as the result of a financial incentive and conditionality mechanism working to align Member State behaviour with the EU's interests. This finding invites a broader question, however, as the literature flags only modest effects of conditional financing in terms of lasting structural reforms – with Dreher (2009) finding that IMF recipient countries overwhelmingly continue to rely on external support over time, and Khemani (2017) arguing that external incentives tend to produce formal compliance without the deeper ownership required for reforms to be sustained. Where does the RRF sit in relation to this pattern? The incentive structure was particularly successful in the policy areas of energy, education, and transport, whilst the EU's residual loss – the CSRs that remain unaddressed – is most pronounced in taxation, pension, housing, and the business environment. Whilst there are signs of the moral hazard problem, such as the 29 disqualified reforms, this concept cannot be further elaborated on from the findings due to the limitations of the research design. The following paragraphs interrogate what the findings can and cannot tell us about whether the RRF represents something genuinely new or a familiar story of conditional financing falling short of its purpose.

The financial incentive made a difference – but drawing lessons from the findings is a cautionary tale. The unstandardised coefficient for the OLS implies that a 1 percentage-point increase in the envelope as a share of GDP in 2023 is associated with a 1.5 percentage-point increase in the share of addressed CSRs in the same Member State. However, does that mean Member States would have addressed more CSRs had they received a larger envelope? The distribution across policy areas suggests the answer is not straightforwardly yes. Some types of CSRs have been less susceptible to RRF reforms than others – and the residual loss concentrated in pension, taxation, housing, and the business environment points to a ceiling on what the financial incentive can deliver, regardless of its size. The composition of policy areas also varies across Member States, meaning that a larger envelope would not necessarily have unlocked more reform activity if the remaining unaddressed CSRs fall precisely in the areas where domestic political constraints or legislative complexity place them beyond the reach of the instrument.

What explains the Member States that perform notably above or below what their envelope size would predict – and what do these outliers reveal about the limits of the analysis? Austria has had 84 pct of its envelope paid out against an EU average of 68 pct, indicating strong implementation capacity and reform commitment that appears to amplify the effect of the incentive – it pursued broad CSR coverage across 28 reforms addressing 15 unique CSRs, including a pension reform. Cyprus is the second smallest economy in 2023, but has a significantly larger envelope than Malta, the smallest (Eurostat, 2023); thus, the envelope is a relatively larger incentive than merely the share of GDP for Cyprus. However, Bulgaria and Estonia both perform below expectations, pointing towards a different set of constraints. Bulgaria has held eight general elections since 2021, a symptom of political instability that might have limited the administrative capacity to implement the RRF plan (Spirova, 2025, p. 64). Estonia's under-performance is due to pressures during the planning phase, according to Eihmanis (2025): the Commission steered Estonia towards social reforms against national preferences, as Estonia is an outlier in this field, resulting in a plan dominated by healthcare and social protection rather than the energy and transport reforms addressed by most other Member States. Estonia's lower score, therefore, reflects the composition of its plan – a distinction the aggregate analysis cannot capture, and a reminder that the variation left unexplained by the regression reflects real differences in planning-phase dynamics and domestic political capacity that the financial incentive alone cannot account for.

Can the study provide preliminary insights into the workings of the RRF instrument in relation to driving structural reforms? The answer is necessarily provisional, as the analysis concludes four months before the deadline for submitting milestones and targets for payment, and according to the Commission, only 65 pct of all milestones and targets related to reforms are labelled as fulfilled on 1 May 2026 (Commission, 2021). However, expectations were that milestones and targets for reforms would be front-loaded to the first years of the RRF, with investment projects reported throughout and towards the end of the programme period due to the time-intensive nature of construction and large-scale initiatives, suggesting that the bulk of reform activity should already be visible in the data (ECA, 2025a, p. 39). Additionally, a sample check of two Member States revealed that many unfulfilled milestones and targets were attached to reforms already registered in this dataset. It is therefore not clear whether the outstanding milestones and targets would lead to a broader coverage of CSRs. With Member States having had four to five years to implement the reforms they planned, the four months remaining are not much, and the picture that emerges from nearly five years of implementation is unlikely to change fundamentally in the time remaining. Whether the unfinished 35 pct of milestones and targets for reforms represents a genuine residual loss from the conditionality structure of the implementation phase, therefore, remains an open question.

Can the EU promote structural reforms in its Member States by rewarding them financially? Keeping the assumptions of the principal-agent framework – that actors are rational and respond to financial incentives – the study finds that the answer is a cautious yes: Member States with larger envelopes relative to GDP addressed a larger share of their CSRs through implemented reforms, and the relationship is statistically significant and robust across multiple specifications. To some degree, the EU can pay its Member States to do structural reforms. However, the study does not assess the degree to which each matched reform actually addresses its corresponding CSR, meaning the finding speaks to coverage rather than depth, and residual loss is present across all Member States and all policy areas even before the question of depth is applied. This tentatively places the RRF closer to the lever end of the spectrum that the broader conditionality literature describes – an instrument that appears to have assisted rather than merely compelled reform activity. There are, nonetheless, signs that the financial incentive enabled national elites to push through reforms of genuine substance – Bulgaria's anti-corruption reform, for instance, established a new anti-corruption body with investigative powers, introduced judicial review of prosecutorial decisions, and transposed EU whistleblower protections, going considerably beyond a minimal response to its recommendation on improving the effectiveness of public administration (CSR Bulgaria, 2020). Whether this reflects the lever dynamic that Jacoby and Hopkin (2019, p. 24) identify as the condition under which conditionality works best, or whether it is the exception rather than the rule, is a question the present study raises but cannot resolve.

A further dimension worth considering is whether the asymmetric financial incentive structure of the RRF implicitly reproduces a creditor-debtor dynamic within the EU itself. The grant component of the RRF – financed through common debt – is repaid collectively by all Member States, whilst the 12 Member States that opted to take up loans are repaying those themselves. The net redistributive effect, therefore, falls primarily on the grant component, where Member States in Southern and Eastern Europe received considerably larger allocations relative to GDP than their Northern and Western counterparts. De la Porte and Jensen (2021, p. 2) find that strict conditionality on the CSRs was central to convincing the sceptical northern Member States to accept the instrument, suggesting the conditionality was designed as much to assure creditors of sufficient oversight as to drive reform. In this light, the positive correlation between envelope size and CSR uptake could be read as a success story – the Member States with the relatively largest envelope addressed a larger share of their recommendations, which is precisely what the creditor Member States demanded as a condition for their agreement. However, this study cannot assess the depth or quality of the reforms behind that coverage, and the risk of the moral hazard problem – that Member States selected reforms that were easy to document rather than those most needed — is most consequential precisely where the grants, and therefore the political expectations, are largest. Whether the conditionality structure has delivered genuine structural change, or a well-documented approximation of it, remains an open question.

6.1 The RRF in the broader picture

Situating the finding within the broader literature on conditionality requires care, because the differences between the RRF and the canonical cases of conditional lending are substantial. The IMF and World Bank conditionality literature documents a persistent pattern of failure, with Dreher (2009) finding that recipient countries overwhelmingly continue to rely on external support over time, and Feldstein (1998) and the Meltzer Commission (2005) arguing that international institutions lack the expertise and legitimacy to impose structural conditions on sovereign states. However, the RRF departs from this model in several important respects. The CSRs are not externally imposed conditions but recommendations that Member States participate in producing through the European Semester cycle – meaning the reform agenda carries a degree of domestic legitimacy (Arnal et al., 2025, p. 1258). The Member States collectively agreed on the RRF instrument, designed their own RRF plans in consultation with the Commission, and are themselves participating in repaying the common debt that finances it (Bekker, 2021). Nevertheless, when studying domestic ownership of the RRF plans, Munta et al. (2024, p. 3787) find that the picture is less straightforward for domestic social actors. The finding that financial incentives produced measurable CSR uptake may therefore say less about conditionality in general than about the specific political and institutional conditions under which the RRF was designed and agreed.

How does the RRF's conditionality structure compare to the typology developed by Jacoby and Hopkin (2019), and does the instrument qualify as a lever in their terms? Their framework identifies three conditions under which conditionality is more likely to succeed: supportive national elites, politically non-controversial policy areas that do not touch core state functions, and EU control of a credible and clearly defined reward (Jacoby & Hopkin, 2019, p. 24). The RRF meets several of these conditions more closely than the Eurozone crisis programmes did(Arnal et al., 2025, pp. 1259–1260). However, the RRF departs from the lever model in one important respect: several of the CSRs the instrument is meant to drive concern precisely the core state functions – pensions, taxation, and labour market policy – where Jacoby and Hopkin find conditionality least effective. The distribution of residual loss across policy areas in this study is consistent with that prediction. Furthermore, Jacoby and Hopkin (2019, p. 5) caution that the apparent success of enlargement conditionality was exceptional and has been overattributed – a warning that applies equally to any optimistic reading of the RRF's moderate positive result. The RRF, therefore, sits closer to the lever end of the spectrum than the club-like Eurozone crisis programmes, but the residual loss concentrated in politically sensitive policy areas suggests that even well-designed ex ante conditionality has its limits.

Is the RRF well-suited to the type of reforms it was designed to incentivise – or does the instrument's measurement logic work against its own ambitions? The RRF is designed in keeping with new public management principles: the idea that structural reforms can be measured, benchmarked, and rewarded on the basis of qualitative or quantitative documentation (Thinus, 2025). This research design is loyal to that output-orientation, relying exclusively on reforms that have been verified and disbursed – and in that sense, the correlation identified in this study is a product of that same logic. However, Efstathiou and Wolff (2018) and Bekker (2021) note that the CSRs are broad, normative, and describe structural change processes that may require many years to materialise – making them an awkward fit for an instrument that measures compliance through verifiable milestones and targets. The European Court of Auditors (2023, p. 4) sharpens this concern, finding that the payment model measures a step-by-step implementation process rather than genuine performance or results. It is perhaps, then, a strange match: a new public management instrument designed to measure output, applied to recommendations that were never intended to be output-measurable in the short- or medium term. This tension does not invalidate the finding as the financial incentive does appear to have moved Member States to address their CSRs – but it raises the question of whether a performance-based instrument measuring discrete outputs is the right tool for driving the kind of broad, lasting structural change the recommendations were designed to produce. As the literature review establishes, neither the political science literature nor the RRF regulation itself offers a commonly agreed definition of what constitutes a reform, meaning the instrument is measuring something that remains conceptually contested.

What do the findings suggest for the future use of the performance-based payment model and financing of reforms in line with the CSRs, as proposed in the 2028–2034 multiannual financial framework? The positive correlation between envelope size and CSR uptake offers some empirical basis for retaining the financial incentive structure, suggesting the model is capable of producing measurable alignment between Member States' reform activity and the EU's recommendations. However, the findings also counsel caution on two fronts. First, with 35 pct of reform-related milestones and targets still unfulfilled at the time of writing, some funding will not find its target, and a meaningful share of planned reforms appears unlikely to be completed before the August 2026 deadline, raising questions about the successor instrument's capacity to manage a potentially even larger programme. Second, the case of Hungary suggests that the performance-based model carries an underappreciated strength beyond its reform-driving function: by conditioning payments on rule of law protections, it proved an effective mechanism for protecting the EU's financial interests despite the reform conditionality failing entirely.

A further lesson from the findings concerns the fit between the instrument and the recommendations it is meant to drive. If the successor budget retains the performance-based payment logic, there may be a case for reformulating the CSRs themselves – making them more specific, measurable, and time-bound – so that the recommendations are better suited to a system that rewards clearly defined and verifiable steps rather than the broad, normative structural change processes they currently describe. It is worth recalling that the RRF was a genuinely novel instrument, conceived under crisis conditions and implemented for the first time without the benefit of prior experience – and some of the weaknesses this study identifies may reflect the inevitable imperfections of a first attempt rather than fundamental design flaws (Arnal et al., 2025, p. 1259). The 2028–2034 multiannual financial framework could be the first opportunity to apply the lessons of that trial-and-error experience deliberately.

7. Conclusion: a cautious yes

This thesis set out to investigate whether the EU can promote structural reforms in its Member States by rewarding them financially, and the answer is a cautious yes. The analysis finds a statistically significant and positive relationship between the relative size of each Member State's Recovery and Resilience Facility envelope and the share of its sub-level country-specific recommendations addressed through implemented reforms, with a Pearson's r of 0.60 and a Spearman's ρ of 0.66, both significant at $p < 0.001$ and robust across multiple OLS specifications. Keeping the assumptions of the principal-agent framework – that actors are rational and respond to financial incentives – the finding supports the proposition that a credible financial reward produces measurable alignment between the agent's behaviour and the principal's goals (Jensen & Meckling, 1976, p. 308). To some degree, the EU can pay its Member States to do structural reforms.

However, the finding speaks to coverage rather than depth. The study matches implemented reforms to the CSRs without assessing how far each reform actually addresses the underlying recommendation – meaning the correlation captures the scope of the EU's reach, not the quality of what that reach delivered. Residual loss, the result of a misalignment between the EU's and the Member States' interests, is present across all Member States and all policy areas, even before assessment of tangible implementation. The distribution across policy areas reveals that the incentive structure was particularly successful in energy, transport, and education, whilst it proved insufficient in pension, taxation, housing, and business environment – areas characterised by high political salience, strong national competence, and limited EU enforcement capacity. This pattern is consistent with both the principal-agent framework's prediction that some divergence between principal and agent will always persist (Jensen & Meckling, 1976, p. 308), and with Jacoby and Hopkin's (2019, p. 24) finding that conditionality is least effective where it touches core state functions, and that CSRs in fields with less existing EU regulation experience lower levels of implementation (Deroose & Griesse, 2014, p. 1). A high uptake in CSRs related to energy is partly attributable to the RepowerEU framework, ringfencing funding to energy policy measures, introduced in response to Russia's invasion of Ukraine. This reflects the principal strengthening the RRF's ambition in this policy area, aligning the interests of the EU and its Member States more closely and resulting in both sides doubling down on energy recommendations.

The RRF departs meaningfully from the canonical cases of conditional lending that the broader literature has found persistently ineffective. Its ex-ante reward structure and performance-based financing logic place it closer to the lever dynamic that Jacoby and Hopkin (2019) identify as the condition under which conditionality succeeds – assisted rather than compelled reform activity, grounded in Member State ownership of the RRF plans. However, the two-phase governance structure complicates this characterisation. Where new governments found themselves bound by their predecessors' commitments, the instrument began to resemble a club. After nearly five years of implementation, 35 pct of reform-related milestones and targets remain unfulfilled, suggesting that the ownership of the lever model depends upon is not always sustained across electoral cycles, as pointed out in the study of the CSRs by Deroose and Griesse (2014, p. 1).

An avenue for future research, besides evaluating the finalised implementation, concerns the relationship between democratic legitimacy, national ownership, and implementation through instruments like the RRF. After nearly five years, 35 pct of reform-related milestones and targets remain unfulfilled, and the democratic legitimacy problem identified in the theory section may be part of the explanation. Where new governments inherited RRF plans negotiated by predecessors and felt limited ownership of the committed reforms, implementation may have stalled compared to Member States with less political upheaval. A case study design selecting Member States that experienced significant government turnover after their plans were adopted and comparing their implementation rates and renegotiation trajectories to those with greater political continuity would be well placed to test whether the democratic legitimacy concern is a genuine driver of residual loss. The findings would carry direct implications for a possible continuation with the novelties of the RRF.

Acknowledgements

I want to thank Casper Ristorp Thomsen for a stimulating discussion on the technicalities of the Recovery and Resilience Facility.

In addition, I am deeply grateful for Jan Smith Høst's feedback on an early version of this work.

Bibliography

Acemoglu, D. (2003). Why not a political Coase theorem? Social conflict, commitment, and politics. *Journal of Comparative Economics*, *31*, 620–652. https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S014759670300101X?via%3Dihub

Alesina, A., & Drazen, A. (1991). Why Are Stabilizations Delayed? *The American Economic Review*, *81*(5), 1170–1188.

Armingeon, K., de la Porte, C., Heins, E., & Sacchi, S. (2022). Voices from the past: economic and political vulnerabilities in the making of next generation EU. *Comparative European Politics*, *20*(2), 144–165. https://doi.org/10.1057/s41295-022-00277-6

Arnal, J., Riesgo, J. P., Niño, I., & Escobar, C. (2025). The Day After NextGenerationEU: What Could the EU Do? *Journal of Common Market Studies*, *64*(3), 1256–1271. https://doi.org/10.1111/JCMS.70034;WGROUP:STRING:PUBLICATION

Aylott, N., & Bolin, N. (2023). A new right: the Swedish parliamentary election of September 2022. *West European Politics*, *46*(5), 1049–1062. https://doi.org/10.1080/01402382.2022.2156199

Basil, O., & Nwankwo, C. (2010). The dialectics of reform: The theory and methodological praxis of reform. *African Journal of Political Science and International Relations*, *4*(5), 181–187. http://www.academicjournals.org/ajpsir

Bekker, S. (2021). The eu's recovery and resilience facility: A next phase in eu socioeconomic governance? *Politics and Governance*, *9*(3), 175–185. https://doi.org/10.17645/pag.v9i3.4290

Busato, F., & Coletta, C. M. (2017). A moral hazard perspective on financial crisis. *Banks and Banks Systems*, *12*(3), 298–307. https://doi.org/10.21511/bbs.12(3-1).2017.13

Capano, G. (2009). Understanding Policy Change as an Epistemological and Theoretical Problem. *Journal of Comparative Policy Analysis: Research and Practice*, *11*(1), 7–31. https://doi.org/10.1080/13876980802648284

Capano, G., & Howlett, M. (2009). Introduction: The Determinants of Policy Change: Advancing the Debate. *Journal of Comparative Policy Analysis: Research and Practice*, *11*(1), 1–5. https://doi.org/10.1080/13876980802648227

Capati, A., & Christiansen, T. (2025). Legitimacy and political dissensus in the implementation of the Recovery and Resilience Facility: the case of Italy. *European Political Science*, *24*(3), 553–577. https://doi.org/10.1057/S41304-025-00538-4

Cerna, L. (2013). *The Nature of Policy Change and Implementation: A Review of Different Theoretical Approaches*. https://www.luciecerna.com/uploads/5/1/2/5/5125288/the_nature_of_policy_change_and_implementation.pdf
Clark, T., Foster, L., Sloan, L., & Bryman, A. (2021a). Quantitative Data Analysis. In T. Clark, L. Foster, L. Sloan, & A. Bryman (Eds.), *Bryman's Social Research Methods* (6th ed., pp. 315–346). Oxford University Press.
Clark, T., Foster, L., Sloan, L., & Bryman, A. (2021b). Using Existing Data. In T. Clark, L. Foster, L. Sloan, & A. Bryman (Eds.), *Bryman's Social Research Methods* (6th ed., pp. 293–314). Oxford University Press.
Commission, E. (2021). *Recovery and Resilience Scoreboard*. European Commission. https://ec.europa.eu/economy_finance/recovery-and-resilience-scoreboard/milestones_and_targets.html
Commission, E. (2024). *Proposal for a COUNCIL IMPLEMENTING DECISION amending Implementing Decision (EU) (ST 7772/22 INIT; ST 7772/22 ADD 1) of 4 May 2022 on the approval of the assessment of the recovery and resilience plan for Sweden*. European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52024PC0557
Commission, E. (2025). *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing the European Fund for economic, social and territorial cohesion, agriculture and rural, fisheries and maritime, prosperity and security for the period 2028-2034 and amen…*. European Commission . https://eur-lex.europa.eu/resource.html?uri=cellar:0d5ded06-639d-11f0-bf4e-01aa75ed71a1.0001.02/DOC_1&format=PDF
COUNCIL RECOMMENDATION of 9 July 2019 on the 2019 National Reform Programme of Czechia and Delivering a Council Opinion on the 2019 Convergence Programme of Czechia, Pub. L. C 301/03, Official Journal of the European Union (2019).
COUNCIL RECOMMENDATION of 20 July 2020 on the 2020 National Reform Programme of Bulgaria and Delivering a Council Opinion on the 2020 Convergence Programme of Bulgaria, Pub. L. C 282/8, Official Journal of the European Union (2020).
COUNCIL RECOMMENDATION of 20 July 2020 on the 2020 National Reform Programme of Hungary and Delivering a Council Opinion on the 2020 Convergence Programme of Hungary, Pub. L. C 282/107, Official Journal of the European Union (2020).
COUNCIL RECOMMENDATION of 20 July 2020 on the 2020 National Reform Programme of Romania and Delivering a Council Opinion on the 2020 Convergence Programme of Romania, Pub. L. C 282/149, Official Journal of the European Union (2020).
Council Recommendation of 21 October 2024 on Economic, Budgetary, Employment and Structural Policies of Denmark, Pub. L. C/2024/6811, Offician Journal of the European Union (2024). http://data.europa.eu/eli/reg/2011/1176/oj
COUNCIL RECOMMENDATION of 9 July 2019 on the 2019 National Reform Programme of the Netherlands and Delivering a Council Opinion on the 2019 Stability Programme of the Netherlands, Pub. L. C 301/112, Official Journal of the European Union (2019).
COUNCIL RECOMMENDATION of 9 July 2019 on the 2019 National Reform Programme of Hungary and Delivering a Council Opinion on the 2019 Programme of Hungary, Pub. L. C 301/101, Official Journal of the European Union (2019).
COUNCIL RECOMMENDATION of 9 July 2019 on the 2019 National Reform Programme of Italy and Delivering a Council Opinion on the 2019 Programme of Italy, Pub. L. C 301/69, Official Journal of the European Union (2019).

COUNCIL RECOMMENDATION of 9 July 2019 on the 2019 National Reform Programme of Spain and Delivering a Council Opinion on the 2019 Programme of Spain, Pub. L. C 301/48, Official Journal of the European Union (2019).

de la Porte, C., & Jensen, M. D. (2021). The next generation EU: An analysis of the dimensions of conflict behind the deal. *Social Policy and Administration*, *55*(2), 388–402. https://doi.org/10.1111/spol.12709

Deroose, S., & Griesse, J. (2014). *Implementing economic reforms-are EU Member States responding to European Semester recommendations? 37*. https://doi.org/10.2765/73044

Dreher, A. (2009). IMF conditionality: Theory and evidence. *Public Choice*, *141*(1–2), 233–267. https://doi.org/10.1007/s11127-009-9486-z

ECA. (2023). *Special report The Recovery and Resilience Facility's performance monitoring framework*.

ECA. (2024). *2023 Annual reports*. http://europa.eu

ECA. (2025a). *Performance-orientation, accountability and transparency-lessons to be learned from the weaknesses of the RRF*. https://www.eca.europa.eu/ECAPublications/RV-2025-02/RV-2025-02_EN.pdf

ECA. (2025b). *RRF support for an improved business environment Only partially addresses country-specific recommendations, but some first results contributed to progress in their implementation*. https://www.eca.europa.eu/ECAPublications/SR-2025-21/SR-2025-21_EN.pdf

Efstathiou, K. ;, & Wolff, G. B. (2018). Is the European Semester effective and useful? *Bruegel Policy Contribution*, *9*. https://hdl.handle.net/10419/208014

Eihmanis, E. (2025). Enforcing social standards through RRF conditionalities: the European Commission as a social advocate in Latvia and Estonia. *Comparative European Politics 2025 23:6*, *23*(6), 1142–1161. https://doi.org/10.1057/S41295-025-00441-8

European Commission. (n.d.). *Country-specific recommendations database*. Retrieved 13 March 2026, from https://ec.europa.eu/economy_finance/country-specific-recommendations-database/

Eurostat. (2023). *Gross domestic product, volumes*. Eurostat Database. https://ec.europa.eu/eurostat/databrowser/view/teina011/default/table?lang=en

Feldstein, M. (1998). Refocusing the IMF | Foreign Affairs. *Foreign Affairs*. https://www.foreignaffairs.com/articles/asia/1998-03-01/refocusing-imf

Ferejohn, J. A. (1986). Incumbent Performance and Electoral Control. *Public Choice*, *50*(1–3), 5–25.

Fernandez, R., & Rodrik, D. (1991). Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty. *The American Economic Review*, *81*(5), 1146–1155.

Friedman, S. D. (2008, November 25). Do Not Waste This Crisis. *Harward Business Review*. https://hbr.org/2008/11/dont-waste-this-crisis

Guidi, M., Piccinetti, M., & Verzichelli, L. (2023). NextGenerationEU and the European Semester: Comparing National Plans and Country-Specific Recommendations. *Politics and Governance*, *11*(4), 339–351. https://doi.org/10.17645/pag.v11i4.7351

Hall, P. A. (1993). *Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain* (Vol. 25, Number 3).

Hooghe, Liesbet., & Marks, Gary. (2001). Multi-level governance in the European Union. In *Multi-Level Governance and European Integration* (pp. 1–32). Rowman & Littlefield Publishers.

Huntington-Klein, N. (2025). Chapter 13 - Regression | The Effect. In N. Huntington-Klein (Ed.), *The Effect: An Introduction to Research Design and Causality* (2nd ed., 13). Chapman & Hall/CRC. https://theeffectbook.net/ch-StatisticalAdjustment.html#ch-StatisticalAdjustment

Jacoby, W., & Hopkin, J. (2019). *From Lever to Club? Conditionality in the European Union During the Financial Crisis*.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Kalyuzhna, N. H., & Shchelkunov, I. V. (2025). The EU Recovery and Resilience Facility as a Tool for Macroeconomic Stabilization. *Problemi Ekonomiki*, *3*(65), 19–27. https://doi.org/10.32983/2222-0712-2025-3-19-27

Khemani, S. (2017). Political Economy of Reform. *World Bank Group Development and Research Group*. http://econ.worldbank.org.

Meltzer, A. (2005). *The Report of the International Financial Institution Advisory Commission: Comments on the Critics*. https://doi.org/10.1184/R1/6708722.V1

Munta, M., Pircher, B., & Bekker, S. (2024). Ownership of national recovery plans: next generation EU and democratic legitimacy. *Journal of European Public Policy*, *31*(11), 3787–3811. https://doi.org/10.1080/13501763.2023.2248202

Radelet, S., & Sachs, J. D. (1998). The East Asian Financial Crisis: Diagnosis, Remedies, Prospects. *Brookings Papers on Economic Activity*, *1*.

Regulation (EU) 2021/241 (2021).

Spirova, M. (2025). Bulgaria: Political Developments and Data in 2024. *European Journal of Political Research Political Data Yearbook*, *64*(1), 64–87. https://doi.org/10.1111/2047-8852.70025;JOURNAL:JOURNAL:20478852;WGROUP:STRING:PUBLICATION

Streeck, W., & Thelen, K. (2005). Introduction: Institutional Change in Advanced Political Economies. In W. Streeck & K. Thelen (Eds.), *Beyond Continuity : Institutional Change in Advanced Political Economies* (pp. 1–39). Oxford University Press.

Thinus, P. (2025). Rule of Law Protection through RRF Performance-based Spending Conditionality: Lessons from the Polish, Hungarian and Slovak Plans. *European Journal of Risk Regulation*, *16*(3), 882–895. https://doi.org/10.1017/ERR.2025.10059

Varga, J., & Veld, J. in 't. (2013). The growth impact of structural reforms. *Quarterly Report on the Euro Area (QREA)*, *12*(4), 17–27. https://ideas.repec.org/a/euf/qreuro/0124-02.html

Wójtowicz, A. (2024). EU Energy Security After Russia's Invasion of Ukraine – Substance, Strategy and Lobbying. *Studia Europejskie*, *28*(2), 157–171. https://doi.org/10.33067/SE.2.2024.8

Zeitlin, J., Bokhorst, D., & Eihmanis, E. (2025). Governing the European Union's recovery and resilience facility: National ownership and performance-based financing in theory and practice. *Regulation and Governance*, *19*(3), 864–884. https://doi.org/10.1111/REGO.12619

Zsiros, S. (2026, May 13). *Hungary's Magyar set to outline his red lines over EU funding talks in a letter to von der Leyen | Euronews*. Euro News. https://www.euronews.com/my-europe/2026/05/13/hungarys-magyar-set-to-outline-his-red-lines-over-eu-funding-talks-in-a-letter-to-von-der-