**RESEARCH ARTICLE** 



# The pursuit of welfare efficiency: when institutional structures turn 'less' into 'more'

Christina Steinbacher<sup>1</sup>

Accepted: 3 May 2024 / Published online: 23 May 2024 © The Author(s) 2024

# Abstract

Addressing current socio-economic crises strains public budgets and may threaten fiscal sustainability. Particularly in the welfare sector, where high expenditures meet poor controllability, efficient resource usage is essential to ensure future governments' capability to act while alleviating current problems. Consequently, this paper asks: why are some countries more efficient in translating social expenditure into welfare outcomes? To answer this question, it is argued that efficiency is a matter of institutional structures and their vertical policy-process integration (VPI): efficiency depends on institutional structures' capability to (1) ensure policymakers' responsibility and to (2) provide coordinated feedback, thus pushing for considerate and informed resource use. Analysing the effect of VPI on the relationship between welfare efforts and social outcomes in 21 OECD countries over three decades, the results show that VPI can not only turn 'less' into 'more', but it also compensates for performance losses in the face of spending cuts.

**Keywords** Government efficiency · Social policy · Welfare state · Vertical policy-process integration · Institutional structures

# Introduction

In view of current socio-economic crises, the role of the state in managing them and in alleviating their consequences moved once again to the centre of public debates. Calls for greater government support are growing louder. Given the general reluctance and resistance to tax increases, more debt is being taken on and budget targets are being postponed or abandoned. It still appears "that 'throwing money at issues' is [...] the first political reflex" (Hauner & Kyobe, 2010: p. 1527). However, these dynamics pose significant risks for the future ability of states to act. Current burdens are being passed on to future generations and the current rise in inflation is making them increasingly difficult to bear.

Welfare states are particularly concerned about these problems, but not only because social safety nets are of great societal relevance. Two reasons make social expenditure the

Christina Steinbacher christina.steinbacher@gsi.uni-muenchen.de

<sup>&</sup>lt;sup>1</sup> Department of Political Science, LMU Munich, Oettingenstraße 67, 80538 Munich, Germany

specific focus of attention: first, the sheer size and considerable growth of social spending relative to other budget items across countries. Despite national differences in welfare approaches, the current per capita spending of many traditional 'skinflints' compares well to that of classic 'spendthrifts' of the not-too-distant past.<sup>1</sup> Second, state-led steering of social expenditure is exceptionally difficult, with spending amounts being largely driven by pre-existing policy entitlements and situational imperatives (Scruggs, 2007). Governments face intense hindrances when trying to alter course due to path dependencies and vested interests (Adam et al., 2019; Merrien, 1998; Powell & Barrientos, 2004).

Thus, the greater the spending and the higher the pressure on public budgets, the more urgent become questions of how to use capital most effectively: how can savings be made without compromising results? Even though several scholars have already highlighted efficiency differences between welfare states (Afonso & Kazemi, 2017; Afonso et al., 2010; Cantillon et al., 2003; Longford & Nicodemo, 2010; Valls-Fonayet et al., 2020), the resulting patterns deviate from specific welfare regime types and necessitate further elucidation. Therefore, this paper aims to address the question: why are some countries more successful in translating their national welfare efforts, specifically social spending, into welfare outcomes than others?

The efficiency of welfare states is determined by the effectiveness of social policy measures in relation to their costs. Improving efficiency entails enhancing policy effectiveness and reducing expenditure. However, coordinating these two factors poses challenges. On one hand, studies have shown a strong correlation between the effectiveness of redistributive measures and the level of transfers or expenditures (Gugushvili & Laenen, 2021). On the other hand, our understanding of how different welfare strategies or combinations of social policy instruments lead to more efficient results is limited and partly inconclusive, given pronounced contextual dependencies in social policymaking (Antonelli & Bonis, 2019; Bressers & Klok, 1998; Valls-Fonayet et al., 2020). As welfare systems change and hybridize, comparative research on welfare efficiency demands more nuanced conceptual tools than the traditional typologies to cope with context dependence and to explain differences in efficiency between countries and over time (Ciccia & Javornik, 2019; Ebbinghaus, 2012; Fernández-i-Marín et al., 2021; Van Kersbergen, 2013).

Against this background, this paper proposes a new comparative perspective that focuses on institutional structures in the social sector rather than on individual measures to explain welfare efficiency. It argues that welfare efficiency depends on the extent to which policymakers are enabled and obligated through institutional mechanisms to pursue social policy options that are cost-effective given their respective social policymaking context. These mechanisms rely on the capacity of politico-administrative arrangements to provide (1) central responsibility for the outcomes of policymaking and (2) the capacity to coordinate and integrate implementation feedback into decision-making. These specific features rendering politico-administrative structures are systematically captured by the concept of vertical policy-process integration (VPI) and its two dimensions, top-down and bottom-up integration (Knill et al., 2021a, 2021b). The mechanisms of top-down responsibility and bottom-up feedback jointly influence policymakers' ability and obligation to use and allocate resources prudently and to pursue efficient welfare solutions. The limitation of resources furthermore amplifies both mechanisms.

<sup>&</sup>lt;sup>1</sup> For a visualization, see Figure A2 in the online appendix.

To test this argument, this paper systematically compares the effect of politico-administrative arrangements on the relationship between welfare efforts and social outcomes in 21 OECD countries over a period of three decades. Across models, the results corroborate considerable effects of VPI on expenditure efficiency, especially when confronted with lower levels of social expenditure or spending cuts. These compensatory effects of VPI may turn 'less' into 'more' and pave the way towards greater efficiency. However, it is also shown that the efficiency effect of VPI has limits: for very high expenditure levels and spending increases, the efficiency effect becomes negligible. These results indicate prerequisites for an effect and reflect the analytical limitations of the study.

Against this background, this research paper offers the first quantitative and comparative evidence on the impact of different politico-administrative structures on the efficiency of public capital employment. It contributes to the literature in three ways. First, it complements research on policy instrument mixes and policy integration by highlighting the importance and precedence of vertically integrated policymaking structures for the quality of policy design choices and effective policy integration. Emphasizing the importance of institutional relationships between implementation and policy design, the paper sheds light on qualities of policymaking systems that receive too little attention in current debates (Wegrich, 2015). Secondly, it offers a nuanced analytical instrument, VPI, to capture and compare the national context of social policymaking beyond traditional welfare state typologies, enriching comparative welfare state literature and partially alleviating problems of context-dependencies in other areas. Thirdly, it provides a comprehensive cross-country and over-time data set on the interplay between social policymaking and policy-implementing institutions, mirroring existing dynamics within welfare states' politico-administrative arrangements (Powell & Barrientos, 2004). This time-dynamic component leads to the final practical contribution: the level of VPI may be modified by reforms. As vertical policy-process integration has been shown to increase the efficiency of social spending, the costs of integrative and coordinative reforms may well be offset by their benefits, allowing for effective control of spending without sacrificing welfare outcomes.

The subsequent parts of the research paper are structured as follows: the paper begins by elaborating on the state of the art and then moves on to the theoretical argument and corresponding hypotheses. This is followed by a section dealing with the general research design and the newly introduced data sets. Subsequently, the main results of the linear panel regression models are presented and discussed. Additional robustness is provided in the next part. Finally, the paper concludes with a discussion of the results and their implications for research and practice.

#### Why the pursuit of welfare efficiency isn't a simple task

The efficiency of a welfare state depends on the effectiveness of its social policies in conjunction with their associated costs. Mere identification of effective policies is insufficient; their cost-benefit ratio must be assessed. Addressing these trade-offs poses challenges for comparative research and permeates the debates in the literature (Bressers & Klok, 1988; Capano & Howlett, 2020; Valls-Fonayet et al., 2020). Three approaches merit particular attention. First, efficiency has been argued to depend on the type of welfare state and the predominant welfare strategy. One of the most prominent debates in this regard revolves around the 'paradox of redistribution' (Korpi & Palme, 1998), questioning the expediency of selective strategies. Korpi and Palme suggest that while targeting may offer short-term efficiency, it jeopardizes long-term support for social policies and weakens the foundation for redistribution. However, Gugushvili and Laenen's (2021) comprehensive literature review demonstrates that twenty-first-century welfare states diverge considerably from expected outcomes: "the only assumption, i.e., unequivocally supported by more recent studies is that higher welfare spending is associated with lower poverty and inequality" (p. 123).

Since the perceived superiority of 'universalism' appears to be primarily rooted in empirical correlations with higher social expenditures (Brady & Bostic, 2015; Jacques & Noël, 2018), and comparative evidence on the associated spending's efficiency is widely contradictory (Antonelli & Bonis, 2019; Valls-Fonayet et al., 2020), the trade-off between costs and benefits of social policies remains unresolved. Instead, the inconclusive evidence underscores the intricate and context-dependent nature of various welfare approaches, such as their governance structures or the mobilization of interests (Brady & Bostic, 2015). Apparently, there is no (re)distributive strategy or 'one-size-fits-all' solution that is generally more efficient (Trubek & Trubek, 2005). While social expenditure is crucial for effective poverty reduction, it remains unclear how different strategies lead to distinctively efficient outcomes.

The second debate revolves around the combination of different policy tools and the 'goodness' of these policy instrument mixes (Capano & Howlett, 2020). Despite pathdependence, policy feedbacks (Pierson, 1993) and lock-in effects that limit the number of alternative choices (Howlett, 2018), modern welfare states show a considerable degree of change in response to challenges and efficiency demands: they hybridize, converge, and adapt (Abou-Chadi & Immergut, 2019; Jensen, 2011b; Powell & Barrientos, 2004), mixing different welfare strategies and social policy tools. Jacques and Noël (2021), for example, find that the combination of universalism with accurate targeting has the potential to "make a more effective use of the state's financial resources" when based on a substantial social budget (p. 27). However, the welfare state toolbox extends beyond (re)distributive policies: regulatory welfare policy complements fiscal redistribution, especially in times of strained social budgets that demand greater 'efficiency' (Levi-Faur, 2014). Even though the extent to which regulatory and fiscal instruments are coupled varies across countries (Trein, 2020), the effectiveness and efficiency of such combinations have not yet been systematically assessed.

Yet, research on policy instrument mixes also grapples with context-dependency in evaluating performance. Since the 'optimality' of a policy instrument mix depends on the coherence of its individual tools and their fit with idiosyncratic governance frameworks (Howlett, 2018), it remains a challenge to compare and generalize the quality of policy choices and their effects (Fernández-i-Marín et al., 2021; Magro & Wilson, 2019). Even supposedly cost-effective strategies, such as market-based instruments (Bakam et al., 2012), have been shown to present suboptimal choices given different circumstances (Steinebach, 2022).

Finally, literature on policy integration argues that the cross-sectoral and multidimensional nature of social policy (Sen, 2006) requires holistic approaches in policy formulation and implementation to reduce uncertainty and conflict (Candel & Biesbroek, 2016; Cejudo & Michel, 2017; Tosun & Lang, 2017). Policy integration is intended not only to guarantee coherence in formulating policy mixes but also to engage relevant stakeholders across sectors and at various government levels in the policymaking process. Integration efforts aim to collectively pursue "a goal that encompasses—but exceeds—the programs' and agencies' individual goals" (Cejudo & Michel, 2017: p. 750).

While showing greater sensitivity to contextual conditions that influence policy integration initiatives (Trein et al., 2021), researchers encounter challenges in delineating the overarching consequences of policy integration. The lack of generalizable evidence results from an overly concentration on single policies and organizations. So far, it remains widely unclear whether the benefits of policy integration can offset or exceed coordinative and integrative costs (Lundin, 2007).

What is more, most policy integration literature primarily addresses cross-sectoral, horizontal fragmentation. Overcoming the vertical fragmentation in policy production and implementation across different levels of government is only rarely made subject of analysis (Homsy et al., 2019; Steinbacher, 2023). However, it has been shown that the procedural integration or 'bureaucratic coupling' between policy formulation and implementation across levels constitutes an integral component of effective policymaking (Fernández-i-Marín et al., 2023) and may alleviate existing implementation problems of policy integration.

In conclusion, previous research on the effectiveness and efficiency of social policies has consistently emphasized their context-specific nature in design and implementation. Since welfare state typologies often struggle to adequately discern modern welfare states' contextual conditions (Arts & Gelissen, 2002; Ebbinghaus, 2012) and for the lack of more fine-grained concepts, research has been prompted to concentrate on case studies on the policy or organizational level. In consequence, a noticeable research gap evolved in the comparative analysis of diverse welfare outcomes across countries and over time. Moreover, it has been shown that so far there are no inherently superior strategies or universal guidelines for achieving welfare efficiency. Instead, the efficiency of welfare states has been suggested to depend on the specific configuration and integration of social policy portfolios, which must navigate different trade-offs and constraints. Yet, also in this regard we find a lack of comparative tools and evidence to substantiate recommendations related to instrument mixes and policy integration.

Against this background, this paper proposes to focus not primarily on policy outputs, but on the underlying institutional structures that condition governments' capability and commitment to identify and pursue policy options that strike a context-dependent 'optimal' balance between costs and effects. In the absence of fundamentally more efficient policies, institutionalized iterative processes are crucial for managing complexity and approximating efficient measures. Yet, institutional structures have so far only played a very limited role in analysing governments' capacity to act (Huber & Stephens, 2001; Merrien, 1998; Scharpf, 2000). Only a few studies show an interest in scrutinizing the effect of at least single qualities of institutional arrangements within welfare states, such as accountability (Malbon et al., 2019), decentralization (Altreiter & Leibetseder, 2015) or bureaucratic quality (Afonso et al., 2010; Cantillon et al., 2003). Sectoral systematizations of institutional set-ups remain largely missing.

# Approaching welfare efficiency through vertical policy-process integration

Since the pursuit of welfare efficiency is not a one-time shot but an iterative and enduring process, this paper argues that its success is conditioned by sectoral institutional structures and how they work together. These structures are captured by the recently introduced concept of VPI (Knill et al., 2021a, 2021b). Even though the concept has been primarily developed for explaining effective policymaking, it will be shown in the following that its logics easily travel and may be applied to efficiency questions. To make the proposed arguments more tangible, insights and anecdotes from interviews with Italian and Irish social policy formulators and implementers complement and illustrate the theoretical considerations. The choice of countries is based on substantial differences in terms of their current VPI and welfare efficiency, approximated by their spending-performance ratio, while facing similar austerity pressures (Bozio et al., 2015). Information on interviews and case selection criteria can be found in the online appendix.

# The concept of vertical policy-process integration

VPI captures the integration and coordination of processes structuring interactions between those who formulate policies and those who implement them in a specific policy field (Knill et al., 2021a). This functional differentiation does not automatically draw lines between different levels of government or institutions but, instead, distinguishes between functional priorities even within entities. When considering national policies, the formulation level is primarily associated, albeit not exclusively, with ministerial bureaucracies interfacing with politics, which typically play a pivotal role in the production and fundamental structuring of policies. In contrast, the implementation level tends to be more diverse and can encompass national, regional, or local actors, as well as decentralized units, responsible for the practical implementation of policies.

To capture the degree of coupling between these two functional 'levels', VPI provides two separate conceptual dimensions, top-down and bottom-up integration, as illustrated in Fig. 1. Top-down integration can be defined as the degree to which the policy formulation level bears responsibility for policy implementation in terms of (1) its formal accountability



Fig. 1 Conceptual considerations of VPI and its two dimensions

and (2) its obligations to provide resources for implementation and (3) to arrange the corresponding organizational structures. Bottom-up integration, in contrast, refers to the extent to which the policy implementation level is involved and integrated into policy design processes. It captures the degree to which implementation structures can provide coordinated feedback via (1) articulation and are permeated by (2) consultation procedures and (3) systematic evaluation mechanisms.

The VPI concept was primarily established to explore and explain the effectiveness of general policymaking. First analyses show that VPI increases effectiveness by matching implementation burdens and available capacities (Fernández-i-Marín et al., 2023). But how does VPI relate to the question of the *efficiency* of *welfare efforts*? First, welfare efforts are direct results of social policy and subsequent budget decisions (Jensen, 2011a; Korpi & Palme, 1998; Wilensky, 1975). Hence, it is policy decisions on regulatory or redistributive measures, processed and influenced by politico-administrative arrangements and their VPI, that constitute the bedrock for theoretical cost-effectiveness. Second, efficiency can be improved by enhancing policy effectiveness and by reducing costs. Whereas the bottom-up channel is primarily expected to increase effectiveness through improved policy design that is in line with complex implementation realities, the top-down channel is argued to be especially relevant for the cost lever by attributing responsibility for outcomes.

#### Top-down integration and the optimization of costs

High top-down integration means that policy formulators bear the responsibility for policies and their outcomes. Accountability ensures that policymakers must expect to be held responsible themselves. They are less able to shift blame if something goes wrong (Hood, 2010). The responsibility of policy formulators for resource and organizational costs of policy implementation is double-edged, especially when policy costs are as high as in social policy: it is not only a matter of possible policy failure or underperformance due to a lack of implementation capacity, but also a matter of potential excessive or wasteful spending (Bonoli et al., 2019). When policy formulation is required to provide the resources and arrange the organizational setup for policy implementation, policymakers have less chances to conceal or shift the costs of policies.

Against this background, top-down integration is argued to compel policymakers to search and opt for effective and economic solutions (Jensen et al., 2014). The responsibility mechanism makes it less likely that they act upon short-term political or electoral interests and make inconsiderate social concessions. Since welfare obligations are usually long-term commitments with potential boomerang effects, policymakers also have an interest in retaining as much control as possible over their political and budgetary decisions, for example by subjecting them to performance checks or sunset clauses. Consequently, the responsibility mechanism not only ensures 'sufficient' administrative capacity within social policymaking but increases the incentives of the policy formulation level to optimize the use of resources. Top-down integration makes the effective use of the cost lever for efficiency more likely.

The Irish response to the financial crisis gives an illustrative example of how high top-down integration affects the employment of social policy resources. Within the Irish Department of Social Protection (DSP), whose upper echelons are concerned with policy formulation whereas its operative sections implement social policies, policy implementers report that implementation structures were centralized to increase efficiency and that a resource request "now [...] has to be signed off by the Assistant Secretary General" (DSP\_1), this is the policy formulation level. Implementers are asked to "do initial trial runs [...] to see if the work [...] is valuable or a waste of time" (DSP\_2). At the same time and despite austerity, all interviewees from the DSP and the Irish Pensions Authority (PA) praise the responsiveness of policy formulation to essential implementation requirements (DSP\_1; \_2; \_3; PA\_1; \_2; \_3). An assistant principal in the PA summarizes: "Our parent department [the DSP] expects us to guide them on resources to meet legislative requirements because we are the ones to know what it's going to take [...] but to be very realistic and conservative about what we need" (PA\_3).

The situation is quite the opposite within Italian social policymaking. The cost lever cannot be effectively utilized due to the policy formulation level's low level of accountability and pronounced negligence towards implementation. In Italy, responsibility for implementation and policy outcomes is widely dispersed between national, regional and local levels; blame is quickly shifted to policy implementers (ROSP\_3; INPS\_2). Implementation requirements are left unattended with fixed budgets and "totally inadequate" staffing (MLSP 2; see also INPS 1; 2; MLSP 1; ROSP 1; 2; 3). Funds may even be closed for ongoing social policy measures: "If we are not quick enough, we run out of money" (INPS\_1). Frequently, legal provisions simply stipulate that "the organization must manage with 'current expenditure'" failing to ensure "organizational sustainability" (MLSP\_1). Ministerial heads of unit confirm that "there is not enough attention on the implementation [...] as if writing something down into law was enough" (MLSP\_2; see also MLSP\_1). Recent policy initiatives were intended to react to this problem aiming at a greater coordination of the implementation level. However, they appear to drive the different actors even further apart due to poor management (INPS\_2; MLSP\_1; \_2; ROSP\_2; \_3). A high civil servant reports for the citizens' income: "There are only three people [in the relevant ministerial unit] to provide indications, clarifications, etc. to 8.000 municipalities, to 20 regions, and to many employees that work on the measure. We added this partnership to manage the relationship with the municipalities and the region [...], not to leave them alone in the implementation of the measure" (MLSP\_2).

#### Bottom-up integration optimizing effectiveness

The bottom-up channel and the associated feedback mechanism are anticipated to enhance efficiency primarily by leveraging the effectiveness rather than the cost lever. Systematic processing of information is decisive for effective policy design (Ansell et al., 2017). In this context, the implementation level assumes a distinctive role as pool and provider of expertise. Even though decisions on (social) policy measures involve compromise and political bickering and do not (and should not) only reflect technocratic considerations, administrative expertise has the potential to soften political conflict and to steer effective solutions reducing the waste of public resources (Polman & Alons, 2021; Ryan, 2001). It is those who implement policies, rather than those who decide them, who right away learn what practical implications different welfare efforts have. It is the implementation level that knows where the money just trickles away and can be saved, or whether transfers are not well-targeted and may be better invested elsewhere (Jacques & Noël, 2021; Skocpol, 1995). They come across conflicts between policies or actors and are aware of side effects (Kern & Howlett, 2009).

Bottom-up integration is argued to ensure that implementation expertise and experience is effectively fed into the policy formulation and calibration processes. Considering the complexity of modern social protection schemes and the variety of actors involved in policy implementation, effective feedback requires horizontal coordination and regular exchange within and between implementation bodies (articulation) to get streamlined and useable feedback (Lundin, 2007; Peters, 2018). Furthermore, the implementation level must be given the opportunity to voice its positions through consultative procedures or sys-

must be given the opportunity to voice its positions through consultative procedures or systematic evaluation of policy implementation. Upon the basis of institutionalized feedback channels, policymakers are able to learn (Sabatier & Mazmanian, 1980) and adapt social programs to context-dependent needs (Ascher, 2023). Especially in social policy, where there is no clear best practice for parsimonious but effective policy solutions (Cammeraat, 2020; Trubek & Trubek, 2005), it is crucial that policymakers receive all available information on what works best when and where (Chindarkar et al., 2017). In this way, bottomup integration ensures that policy goals and instruments align with complex implementation realities, and that resources can be effectively allocated.

Looking at the proposed mechanism in practice, Irish social policy administrators across levels highlight the receptiveness of the policy formulation level and commend very good communication (DSP\_1; \_2; \_3; PA\_1; \_2; \_3). Referring to the DSP's policy design section, implementers report that "they talk to the people on the floor [...] and they do listen" (DSP\_2), that implementers "can speak freely about things to them" (PA\_3), that recommendations and concerns "feed back up the line quite quickly" (DSP\_3), and that "they act on the feedback" (DSP\_2). In the context of joint working groups that integrate actors from different levels and institutions (DSP\_1; \_2; PA\_1), a case officer recounts that he pointed at inconsistencies in a policy's design that were consequently resolved: "Had I not been involved in that workshop or anybody that had my kind of background knowledge, it probably would have went through and would have cost the Department [of Social Protection] money" (DSP\_2).

In Italy, on the other hand, implementers complain about inefficiencies of processes and policies and suggest concrete improvements that would plausibly enhance the effectiveness of individual measures and procedures (INPS\_1; \_2; ROSP\_2; \_3). Their remarks range from optimizing timescales and workflows to identifying loopholes, ineffective resource allocation and capacity requirements. However, they do not see a chance to voice their feedback. "It is then up to us to try to recover the gap between what the State requires and what we can do" (ROSP\_2; see also INPS\_1; \_2; ROSP\_3). Even top ministry officials criticize the fact that they "were never consulted" on certain policies even though these policies directly fell within their remit (MLSP\_2). All respondents consider the ignorant or even antagonistic relationship between policy formulation and implementation a major reason for inefficiencies in the design and rollout of social policies that could be avoided (INPS\_1; \_2; ROSP\_1; \_2; \_3; MLSP\_1; \_2).

#### Expected joint effects of VPI on welfare efficiency

Through these complementary mechanisms of coordinated feedback and responsibility, VPI is expected to moderate the relationship between welfare efforts and social outcomes, increasing the efficiency with which these efforts translate into performance. While top-down integration is assumed to render the volume of welfare efforts as parsimonious as possible, bottom-up integration is focused on the effectiveness of the resource allocation of underlying regulatory and redistributive measures given complex implementation realities. With VPI enabling and compelling social policymakers in the pursuit of welfare efficiency, the first hypothesis reads as follows: *the more robust the coupling of social policy design* 

and implementation through VPI, the greater the efficiency with which social spending translates into positive welfare outcomes.

However, the expected moderator effect of VPI is not assumed to be uniform across different levels of welfare efforts but to intensify with relative and absolute resource shortage for two reasons. First, various studies show that the efficiency of relative spending decreases the bigger the amount of employed capital becomes (e.g., Afonso & Kazemi, 2017; Hauner & Kyobe, 2010). According to diminishing marginal utility and distributive efficiency, efficiency is highest when welfare efforts are received by those with the greatest need. It follows that the marginal utility of welfare efforts must diminish with every coin spent even if the money is 'optimally' allocated through high levels of VPI. Marginal utility and, hence, efficiency of welfare efforts is not only diminishing on the individual recipient level. It also shifts the baseline with each welfare effect achieved. The 'better' people are off, the 'harder' it gets to make a difference through additional social spending.

Second, also feedback and responsibility mechanisms are sensitive to scarcity and intensify their efficacy when confronted with limited or diminishing welfare efforts. Regarding bottom-up integration and the associated feedback mechanism, the value of implementers' expertise and experience for policy design increases. The scarcer the resources spent, the more valuable the information becomes regarding their optimal allocation, as there is less room for error in achieving desired outcomes. Top-down integration further intensifies this link. The scarcer the resources, the stronger the commitment of policymakers to find parsimonious solutions. Competition for resources becomes fiercer and justification pressures on their usage increase. The risk of excessive or wasteful spending shrinks.

This means that VPI's efficiency effect is stronger, the scarcer the welfare efforts. In consequence, the second hypothesis reads as follows: *the scarcer the resources, the stronger the efficiency effect of VPI on the relationship between social spending and welfare outcomes.* 

# Research design

To test how the expected moderator effect of VPI holds up against empirics, I estimate different time-series cross-sectional models with interaction term based on fixed effects and first-differenced estimators. Relying on differently specified models, the robustness of the findings is widely hedged against the difficulties scholars face when working with panel data, such as autocorrelation, unit heterogeneity and non-stationarities.<sup>2</sup>

The main linear panel regression model includes country as well as year fixed effects using the within transformation (Baltagi, 2021) as well as panel-corrected robust standard errors (Bailey & Katz, 2011). It examines the effect of institutional structures on the relationship between welfare efforts and outcomes at their absolute levels. The second model, in contrast, relies on first-differenced estimators. Rather than focusing on absolute levels of welfare efforts and outcomes, it examines their evolution and changes from one year to the next. This approach models shortages and increases of input factors, further elucidating how the VPI effect manifests under varying resource conditions. Thus, the first-differenced model not only ensures compliance with conservative standards by addressing issues of non-stationarity (Engle & Granger, 1987) but also offers an additional perspective on how

<sup>&</sup>lt;sup>2</sup> For a similar treatment of TSCS-data for explaining policy performance, see, Knill et al., (2012).

institutional structures influence the translation of inputs into outcomes. For this second model, country fixed effects are dropped as cross-country idiosyncratic effects largely vanish through the first-differencing procedure. For the same reason, those control variables are removed whose inclusion is primarily based on cross-country differences but not on within-country changes (Kittel & Winner, 2005).<sup>3</sup>

Finally, in terms of data treatment, all independent variables are standardized to a standard deviation of one for better comparison. Yet, it has been refrained from rescaling them to a mean of zero: first, the original data set only includes positive values, and second, fabricated negative values may confuse the first-differenced and respective interaction terms. Furthermore, and as common in the literature, all independent 'input' variables are lagged by one year. All independent 'institutional' variables, in contrast, are lagged by two years for estimating realistic effects on performance.

Upon this basis, the analyses track the evolution of national social spending and the level of VPI in their interactive relationship with social policy performance within a broad sample of 21 OECD countries. This way, high variance was ensured with respect to the core variables. The selected countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and the United States. The observation period stretches from 1990 to 2019. However, not all units are fully observed throughout the period of investigation resulting in a slightly unbalanced panel including 538 observations for the fixed effects model and 517 for the first-differenced model.

#### Measuring welfare outcomes

The central aim of social policy, as perceived here, is to diminish poverty and mitigate the risk of falling into poverty. The extent to which welfare systems effectively achieve this objective serves as the cornerstone for defining and measuring welfare performance in this paper. Nevertheless, identifying an appropriate metric that aligns with this conceptualization of welfare performance and corresponds to the scope of this study presents inherent challenges.

Existing studies often rely on individual indicators which can only partially capture the complex and multidimensional nature of welfare outcomes. Poverty is not solely defined by a lack of income; it also necessitates consideration of individuals' deprivation of 'capabilities' (Sen, 2006). Even though comparative research often utilizes either relative or absolute poverty rates, or measures of income inequality, these measures overly focus on income or on specific welfare approaches (Arts & Gelissen, 2002; Atkinson, 2019; Castles & Mitchell, 1992; Kautto, 2002). Despite their presumed simplicity, individual poverty measures frequently also struggle with comparative validity (Thorbecke, 2004). Consequently, to adequately capture welfare outcomes, an aggregate metric is required.

However, established aggregate measures suitable for analysing welfare outcomes in OECD countries over the last thirty years are hardly available (Greve, 2017; Hagerty et al., 2001; Ranis et al., 2006). Many indices are specifically tailored or even restricted to the developing world (e.g., the Multidimensional Poverty Index). Other indices are available for only a very limited set of years or exclusively rely on subjective data (Greve, 2017). Among the more widely used and broadly available objective indices, the Human

<sup>&</sup>lt;sup>3</sup> Their inclusion, however, does not alter the results; it only slightly decreases the model fit.

Development Index (HDI) stands out as the most prominent example. However, the HDI grapples with insensitivities towards several factors, such as inequality or unemployment, struggling to discern nuances among advanced democracies (Biggeri & Mauro, 2018; Das-gupta & Weale, 1992; Metzger & Shenai, 2022). Aguña and Kovacevic (2010) demonstrate that "the income index is the most significant driver of differences in the HDI" (p. 6), especially in countries with higher degrees of development.

Against this background, this paper suggests a new Welfare Performance Index (WPI) that integrates measures of poverty incidence and poverty risks. In doing so, the WPI also captures substantive structural problems and addresses the deprivation of capabilities that can plausibly be expected to be tackled by welfare states. The WPI conceptually links welfare outcomes to government action. By combining these indicators, the WPI further seeks to overcome the limitations associated with relying solely on poverty rates. These limitations include conceptual narrowness, restricted comparability due to their reliance on national median incomes, and susceptibility to cyclical effects (Clasen et al., 2007; Sen, 2006). Furthermore, the WPI distinguishes between different subfields of welfare performance, ensuring its adaptability and precision.

The WPI consists of three subindices: unemployment, old-age pensions, and family benefits. Each subindex contains two components. The first element comprises the *general welfare outcome*, i.e., the poverty rate among the target group, e.g., the unemployed. It refers to the percentage of people within the defined population who have less than half of the median household income of the same population at their disposal after taxes and transfers. The second element, in contrast, comprises a structural *problem identifier*, e.g., long-term unemployment, representing the risk of falling into poverty. Here, 'structural' is to be understood as opposed to 'cyclical'. The *problem identifier*, hence, tries to include a representative structural socio-economic problem whose solution is primarily based on public policy intervention. The resulting subindices and their composition are described in Table 1.

The aggregation of the index follows a simple additive logic divided by the number of items, whereby its components are sample-standardized and carry equal weights. Ultimately, the overall WPI as well as its subindices range from 0 to 1 with higher values indicating 'better' performance. The evolution of WPI across countries and welfare state types is shown in Figure A1 in the online appendix.

To corroborate the validity of the WPI, I assess its correlation with the established HDI. To mitigate the HDI's limitations, I employ its inequality-adjusted variant, which exhibits heightened sensitivity to distinctions among developed countries and better captures the nuances of complex welfare outcomes (Foster et al., 2005). Despite the limited number of observations, the correlation stands at 0.6, signifying substantial agreement. To address potential time frame effects, the correlation is recalculated for the HDI, resulting in a correlation of 0.49. The value increases when adjusting for cases whose HDI ranking is overly influenced by the income component (UNDP, 2023).<sup>4</sup>

<sup>&</sup>lt;sup>4</sup> Figure A8 in the online appendix shows WPI's correlations with HDI. In the robustness section, moreover, analyses are reiterated using the HDI as an alternative dependent variable to fortify the results against concerns about the WPI's validity.

Table 1 Components of the W	/elfare Performance Index (WPI) a	nd its subindices			
	Subindex	Variable	Dimension	Description	Source
Welfare Performance Index (WPI)	Employment Performance Index	Working population poverty rate	General outcome	Sample: Working age group 18–65	OECD (2021a): IDD
				Measure: The poverty thresh- old is set at 50% of median disposable income in each country after taxes and transfers	
		Working population long-term unemployment rate	Problem identifier	Sample: Working age group 18–65 Measure: Percentage of the labour force currently unemployed for more than 1 year	OECD (2021b): LFS; OECD (2021c): ALFS
	Old-Age Pension Performance Index	Old-age poverty rate	General outcome	Sample: Age group> 65 Measure: The poverty thresh- old is set at 50% of median disposable income in each country after taxes and transfers	OECD (2021a): IDD
		Old-age income distribution	Problem identifier	Sample: Age group>65 Measure: Gini coefficient	OECD (2021a): IDD
	Family Benefits Performance Index	Child poverty rate	General outcome	Sample: Age group 0–17 Measure: The poverty thresh- old is set at 50% of median disposable income in each country after taxes and transfers	OECD (2021a): IDD
		Youth 'not doing anything'	Problem identifier	Sample: Age group 15–19 Measure: Percentage of the youth currently neither employed nor in education or training	OECD (2021d): Family Database

#### Measuring welfare efforts

Welfare efforts are expressed by public social spending measured in US dollars per capita and derived from the OECD's Social Expenditure Database (SOCX) (OECD, 2022). Even though there are lively debates on how to best define, aggregate and compare social expenditure information, the SOCX is considered one of the most comprehensive assessments of countries' social spending (Castles, 2004; De Deken & Kittel, 2007). The SOCX dataset is widely used and available for large time periods facilitating comparative longitudinal studies.

The reliance on social spending is not without controversy: social expenditure does not account for resource allocation, nor is it a direct result of policymaking (Green-Pedersen, 2004). Yet, total expenditure is undoubtedly a defining feature of welfare efforts, without which little can be said about welfare outcomes (Jensen, 2011a; Korpi & Palme, 1998). Ultimately, it depends on the research questions whether social spending is an appropriate measure. An analysis of efficiency across diverse welfare states in conjunction with politico-administrative structures necessitates such a general proxy measured at high levels of aggregation.

The drawbacks of this choice, however, need to be addressed and the robustness of the results secured. In this context, the main concern relates to the constrained controllability of spending levels due to path dependencies and situational imperatives (Clasen et al., 2007; Merrien, 1998; Powell & Barrientos, 2004), meaning that social spending and welfare outcomes are in parts reciprocally linked. To mitigate this issue in the main models, the outcome variable is largely detached from cyclical effects by including structural poverty risks. Moreover, the variables are lagged and supplemented by either fixed effects or first-differenced estimators. Finally, the analysis is also repeated based on replacement rates retrieved from the Comparative Welfare Entitlements Data Set (Scruggs et al. 2013). However, since the level of replacement rates gives no information on overall resource scarcity, only VPI's general efficiency effect can be tested.

#### Measuring vertical policy-process integration

The data collection on the evolution of welfare-specific VPI follows the two-dimensional VPI measurement scheme provided by Knill et al. (2021b).<sup>5</sup> Coding decisions are derived from systematically comparing and analysing secondary literature, in conjunction with official documents on legislation, administration, and reform processes in social policy. Insufficient or conflicting information was clarified through interviews with experts on social policy, public administration, and management. To enhance the validity and consistency of coding practices, each coding decision was cross-referenced and verified against at least two cases, encompassing both an identical and a divergent indicator evaluation.

For an accurate aggregation of the VPI, I assess the indicators' latent dimensions as suggested by Treier and Jackman (2008). Analysing the latent dimensions of the indicators allows to uncover and understand the hidden structures that contribute to differences in values and patterns of change. It provides a more nuanced view on the evolution of institutional structures in social policymaking and facilitates a more accurate

<sup>&</sup>lt;sup>5</sup> Section A in the online appendix describes the concept formation and coding procedure. Tables A1 and A2 give an overview of the indicators' measurement.

and insightful analysis (Pemstein et al., 2010). I perform cutpoint analyses estimating continuous ordinal coordinates to determine the meaning of indicator changes against the assumption of a linear scale. The obtained loadings form the basis for the modelled VPI scores. The final bottom-up and top-down scores range from a minimum of zero to a maximum of one, resulting in an overall VPI score between zero and two.

The visual inspection of VPI over time and across different welfare states in Fig. 2 reveals two important aspects: first, the vertical integration of politico-administrative is by no means static. Significant changes can be observed in almost all countries, most frequently in the period of the 1990s and early 2000s, possibly reflecting the changes in 'governance' brought about by New Public Management (Merrien, 1998). Although VPI 'improvements' seem to be more common, there are also instances of 'deteriorations' in VPI. Second, patterns of VPI differ not only across different types of welfare states, but also significantly within these groups. Types of welfare states display only few commonalities. Liberal welfare systems, for example, slightly tend toward top-down integration, while bottom-up integration is more dominant in social-democratic welfare states. However, these common tendencies hardly compare with the marked variance within the groups.

These observations highlight the potential of VPI for refining existing conceptions of welfare states in two regards: first, with respect to the institutional structures that form the backbone of welfare states but have not yet received sufficient attention; and second, with respect to the ongoing development and dynamization of welfare regimes.



Fig. 2 Development of the VPI of social institutional structures in 21 OECD countries between 1980 and 2020 across types of welfare states (Esping-Andersen, 1990)

#### Alternative explanations

To analyse the effect of social expenditure in relation to VPI on policy performance, I control for several alternative explanations. All models include the evolution of countries' per capita GDP as a standard economic control variable which has frequently been associated with government performance and contours countries' financial capabilities (Metzger & Shenai, 2022). It is also controlled for private social spending as an alternative to public welfare efforts. In addition, the level of adult education, the size of the international migrant population, and the demographic proportion of countries' elderly population are assumed to affect welfare state performance (Antonelli & De Bonis, 2019; Besharov & Call, 2009; Huber & Stephens, 2001).

Furthermore, the model controls for underlying policy design features that may affect welfare performance. Accounting for the arguments of comparative welfare state literature that different welfare strategies lead to different welfare outcomes, I include a variable assessing the respective countries' lenience towards universalistic or means-tested redistribution strategies over time (Coppedge et al., 2019). Furthermore, I include a measure on portfolio composition to account for the effects of different policy mixes on welfare outcomes. The variable captures the proportion of fiscal (e.g., universal allowance) and non-fiscal, regulatory policies (e.g., retention periods) within national social policy portfolios (Adam et al., 2019; Fernández-i-Marín, 2019).

To clearly carve out the individual effect of VPI, the models also consider the effects of the general political-institutional environment, interest intermediation and government capacity of the countries under study. The first aspect is approximated by two control variables: Henisz's (2000) political constraints indicator captures the presence and preference configuration of institutional veto points that may influence the scale and nature of welfare efforts. Meanwhile, the regional authority index (Hooghe et al., 2016) accounts for the decentralization of powers within multilayered policymaking systems, potentially impacting welfare efforts and their outcomes through additional subnational policies. Jahn's (2016) corporatism index furthermore captures the evolution of national systems of interest intermediation over time, while Hanson and Sigman's (2021) state capacity index is included to control for general government capacity.

#### **Results and discussion**

The analysis first confirms that the volume of welfare efforts clearly matters for welfare outcomes. All models demonstrate a positive and statistically significant relationship between welfare inputs, this is financial resources flowing into welfare regimes, and welfare outcomes, this is the performance of welfare states in reducing and preventing poverty. The positive effect of social spending also remains robust when models are fitted with first-differenced estimators. This implies that performance is not only dependent on absolute levels of social spending but that enhanced outcomes also result from relative increases in spending. In consequence, the models confirm a decisive role of welfare efforts in determining the effectiveness of welfare regimes.

But what about their efficiency? Including an interaction term between social spending and VPI, it is tested whether the strength of the relationship between welfare efforts and performance is conditioned by politico-administrative arrangements integrating policy design and implementation. The results confirm such a moderating role: the degree to which social policy formulation and implementation are coupled influences the translatability of welfare spending into outcomes. Besides its considerable size, the interaction effect also proves to be statistically significant and remains robust in the first-differenced models. Figure 3 illustrates how the obtained estimates perform and compare across the different main models.

Hence, VPI moderates the relationship between social spending and welfare outcomes, both in terms of absolute levels and relative changes. Yet, the sign of the interaction coefficient is negative for both models challenging the general positive efficiency effect postulated in hypothesis one. Therefore, the interaction effect is visualized to capture the exact interdependencies between social spending, vertical integration, and welfare outcomes and to determine the extent to which a general efficiency effect of VPI holds and is shaped by the level and evolution of social spending as envisioned by the second hypothesis.

For the fixed effects model, Fig. 4 shows the predicted values of welfare performance as a function of the moderator variable VPI on the one hand, and the main explanatory variable, the absolute level of social spending per capita, on the other. It demonstrates that an unconditional positive efficiency effect of VPI, as envisioned in hypothesis one, cannot be fully supported based on the present model. The VPI effect has clear limits since it becomes negligible for very high levels of social expenditure. The reinforcing effect of 'scarcity', as suggested by hypothesis two, appears not only to amplify the impact of VPI mechanisms but to determine the very existence of a positive efficiency effect.



Fig. 3 Comparison of model estimates based on fixed effects and first-differenced estimators. *Note*: Whiskers indicate 95% credible intervals based on panel-corrected robust standard errors. All variables have been standardized to one standard deviation. The variables 'adult education', 'elderly population', 'regional authority', 'political constraints' and 'corporatism' are dropped from the first-differenced model. The tabular results can be found in Table A4 in the online appendix



**Fig. 4** Predicted values of absolute Welfare Performance depending on VPI at different absolute levels of Social Spending per capita. *Note:* Shaded zones indicate 95% confidence intervals. The histogram at the bottom of the figure shows the distribution of the moderator variable

However, for lower or moderate social budgets, the magnitude of the VPI effect is considerable: small social 'purses' supported by high degrees of VPI may even outperform substantially bigger social 'purses' not backed by vertically integrated institutional structures. The lower the level of social expenditure, the more pronounced the efficiency boost resulting from high VPI. Hence, VPI compensates lower amounts of social spending, but its positive effect seems to disappear at very high levels of social spending.

Two reasons may explain this observation. The usefulness of VPI and especially bottom-up integration may depend on the underlying welfare strategies that determine the complexity of implementation realities and influence spending levels. Universal policies have been shown to come with higher levels of social expenditure (Jacques & Noël, 2018). This finding is supported by the here used data on social spending per capita: Scandinavian 'universal' welfare states lead the expenditure ranking (see Fig. A2 in the online appendix). The implementation of universal social policies, however, can be reasonably assumed to be less demanding and error-prone than the realization of targeted measures, that may imply complex tasks, such as income assessments, and involve diverse administrative actors. If implementation is hence less challenging, the relevance of VPI and especially implementation feedback diminishes. This leads to the second plausible explanation: the coupling of policy design and implementation through VPI incurs costs (Lundin, 2007), and ties up resources, for example, for coordinative activities. Hence, if the utility of vertical integration is simultaneously challenged from two different sides, this is the ample supply of welfare efforts as well as 'simple' implementation standards, the costs of VPI may exceed the achieved benefits.

In contrast to the fixed effects model, the first-differenced model investigates how *changes* in the independent variables' values from one year to the next affect the annual *delta* in welfare performance. While further fortifying the results against autocorrelation, first differences allow for a clear distinction between increases and decreases in the social budget as well as between gains and losses in welfare performance, providing another perspective on the scarcity condition independent from absolute expenditure.

Focusing on the development instead of the absolute level of social spending, the visual inspection of the moderator effect in the first-differenced model in Fig. 5 confirms a pronounced compensatory efficiency effect of VPI when welfare regimes are confronted with retrenchment. Under conditions of high VPI, the negative consequences of spending cuts can be fully offset and may even have the potential to become positive. A robust coupling of policy design and implementation appears to allow for more efficient welfare states in which spending can be curtailed without sacrificing results.

However, Fig. 5 also shows that increases in social expenditure evenly improve welfare outcomes no matter the degree of vertical integration. This observation challenges the proposed theoretical mechanisms that would have expected an observable even though smaller efficiency effect of VPI on spending increases. These results once again raise the question of whether scarcity—represented here by decreases in social spending—determines not only the strength of the efficiency effect, but also its existence.

While certainly highlighting the need for further and more fine-grained studies on the preconditions and effects of vertical integration, I argue that this observation does *not* necessarily undermine the theoretical expectations if we look at the origins of spending increases and the logics of top-down integration. For spending increases, two different and opposed developments are at play that may blur potential efficiency effects of VPI. Increases in social expenditure have been shown to be positively associated with increased debt and to be predominantly driven by (automatic) reactions to socioeconomic strains that amplify the number of entitled persons, while fiscal balances reflecting the responsibility mechanism of top-down integration—enter in negatively and reduce expenditure (Haelg et al., 2022). High levels of top-down integration prevent increases by encouraging governmental responses to rearrange resources and keep spending and access to entitlements controllable. Low levels of top-down integration, in contrast, come with fewer barriers to expenditure increases which in turn affect welfare outcomes. Hence, looking solely at increases in social expenditure, we may be unable



Fig. 5 Expected development of welfare performance in relation to VPI based on the average marginal effect of increases and decreases in social spending. *Note*: Shaded zones indicate 95% confidence intervals. Deltas ( $\Delta$ ) indicate first-differenced estimates, i.e., changes from one year to the next.  $\Delta$  Social Spending is divided into two subsets delineating positive from negative values

to observe an efficiency effect of VPI since it is blanketed by the performance effects of spending.

Against this background, both models provide robust evidence on considerable efficiency effects of VPI when welfare states are based on limited social budgets or are confronted with retrenchment. In consequence, the second hypothesis can be largely confirmed. However, the provided evidence is less robust when it is about the general efficiency effect of VPI as postulated in the first hypothesis. It can neither be fully confirmed nor invalidated. Whether the modelled scarcity acts as a condition cannot be clearly demonstrated either. This is because the analyses are limited in their ability to distinguish between different bases of spending decisions: First, spending decisions can be rooted in different primal welfare strategies that lead to differently complex implementation requirements, influencing the expediency of VPI. Second, increases in expenditure can be the result of uncontrolled increases in beneficiaries due to socio-economic or demographic factors or the result of new or adapted policies. Since VPI is argued to prevent the former scenario and to condition efficiency in the latter, VPI and expenditure effects overlap when no distinction can be made. As corroborated by this study, social expenditure continues to constitute a pivotal element in the functionality of high-performing welfare states. Higher spending levels and increasing expenditure are consistently associated with improved welfare outcomes.

# Robustness models

To explore whether these limitations endanger the assumed general efficiency effect of VPI, the analyses are repeated using an alternative measure for welfare efforts that is less susceptible to confounding logics, specifically addressing the first hypothesis. For this purpose, the main explanatory variable, social spending, is substituted by replacement rates (Scruggs et al., 2013). Using replacement rates not only alleviates endogeneity concerns and shields the independent variable from automatic reactions to the socio-economic situation, but also removes sensitivities with regards to underlying resource limitations and intricacies of underlying welfare strategies. As benefit calibrations, replacement rates present a relative measure. Therefore, diminishing marginal utility, prerequisites for distributive efficiency, or linkages with specific implementation requirements cannot be inferred.

The results of the corresponding models are detailed in Section F of the online appendix. They support the existence of a general positive efficiency effect of VPI when potential confounding logics are removed. Higher and increasing replacement rates consistently contribute to improved welfare outcomes. This relationship is significantly amplified by VPI in the fixed-effects model, as illustrated in Fig. A6. Even though the interaction effect remains statistically insignificant in the first-differenced model, its positive coefficient also suggests a reinforcement of the positive relationship between replacement rates and welfare performance.

To also resolve doubts on the degree to which the results of the main models may be exclusively driven by the novel and not yet established Welfare Performance Index, the models are furthermore rerun using the HDI as an outcome measure. As argued before, the HDI comes with several shortcomings itself and requires the renouncement of economic and educational control variables. However, the results that are detailed in *Section G* of the online appendix, largely confirm the findings of the main models and suggest a statistically significant and negative interaction effect. Figures A10 and A11 illustrate that efficiency

effects are largest when dealing with smaller or decreasing amounts of social expenditure, while becoming inconclusive in the context of very high levels of social spending or expenditure increases.

# Conclusion

Given today's socio-economic crises and strained national budgets, this paper started with the question of how welfare efficiency can be explained. Why are some countries better able to translate welfare efforts into outcomes? Given the current inflation and mounting burdens on future generations, new strategies and avenues to economize wisely and achieve greater impact at lower cost are urgently needed.

Since previous studies struggled with high levels of contextuality, this paper contends that welfare efficiency is a matter of politico-administrative arrangements and their capability to navigate governments through complex trade-off decisions towards greater welfare efficiency. These capabilities are captured by the concept of VPI. Since efficiency can be affected by leveraging costs and effectiveness, it has been argued that the efficiency of welfare efforts depends on the extent to which policymakers are obliged and enabled to pursue efficient welfare solutions through the mechanisms of top-down responsibility and bottom-up feedback. Scarcity, moreover, amplifies the effect of both mechanisms.

The analyses widely confirm that higher levels of VPI may indeed turn 'less' into 'more'. Smaller or decreasing amounts of social expenditure can be fully compensated by vertically integrated institutional structures. However, the efficiency effect of VPI appears to reach its limits when considering very high levels of expenditure and increases in social spending. To explain these observations, it has been argued that the utility of VPI may not only be affected by lower resource constraints but may also be diminished by undemanding implementation requirements. When both factors come together, efficiency gains may disappear. Moreover, since the results confirm the role of social spending as primary driver of welfare performance, the main models suffer from an analytical limitation: they are not able to distinguish between passive and active expenditure increases. Efficiency gains through VPI may remain unobserved when they are blanketed by improved outcomes due to passive spending increases. Despite these limitations with regards to a potential conditionality of the VPI effect at its upper bounds, the evidence provided on the efficiency effect of VPI is far from trivial. In line with the main purpose of this study, the findings highlight and explain how welfare states can achieve high performance with lower funding levels.

In consequence, these findings bear important implications for comparative research on welfare states and social policy, but also for scholars and practitioners concerned with governmental quality and efficiency more generally. The analyses indicate that sectoral institutional structures, and more precisely the coupling of policy design and implementation, play a crucial role in policymaking outcomes. They shed light on politico-administrative factors that remain frequently overlooked in current debates on the determinants of policy outcomes despite their potential to account for complexity and contextuality in policymaking. This way, they provide a new impetus to research on policy instrument mixes and policy integration. However, the findings also indicate that integrative and coordinative solutions may have limits. The potential conditionality of VPI's utility and its two dimensions provide promising avenues for future research. It also remains to be shown how VPI and its individual dimensions influence the composition of policy portfolios and interact with different policy mixes. Moreover, the concept provides a nuanced and time-variant analytical tool to understand and compare the contextuality of social policymaking beyond conventional welfare state typologies that may enrich comparative welfare state research.

Finally, the findings of this study provide empirical evidence that encourages the reform of politico-administrative set-ups towards greater vertical integration to achieve more efficient and sustainable governance. While the modification of responsibility structures may be challenging, an intensification of feedback opportunities for the implementation level is less demanding. Significant improvements may already be achieved through joint working groups or by employing new technologies to streamline and bundle rapid feedback from implementation. Since higher levels of VPI hold considerable efficiency gains in the case of tight social budgets and savings measures, the costs of integrative and coordinative measures and reforms may well be offset by their benefits.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11077-024-09535-8.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was supported by the European Research Council (ACCUPOL Project, Grant. No. 788941).

# Declarations

Conflict of interest The author reports there are no competing interests to declare.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

# References

- Abou-Chadi, T., & Immergut, E. M. (2019). Recalibrating social protection: Electoral competition and the new partisan politics of the welfare state. *European Journal of Political Research*, 58(2), 697–719.
- Adam, C., Hurka, S., Knill, C., & Steinebach, Y. (2019). Policy accumulation and the democratic responsiveness trap. Cambridge University Press.
- Afonso, A., & Kazemi, M. (2017). Assessing public spending efficiency in 20 OECD countries. In B. Bökemeier & A. Greiner (Eds.), *Inequality and finance in macrodynamics* (pp. 7–42). Springer.
- Afonso, A., Schuknecht, L., & Tanzi, V. (2010). Income distribution determinants and public spending efficiency. *The Journal of Economic Inequality*, 8(3), 367–389.
- Aguña, C. G., & Kovacevic, M. (2010). Uncertainty and sensitivity analysis of the human development index. *Human Development Research Paper*, 47(2010), 2010.
- Altreiter, C., & Leibetseder, B. (2015). Constructing inequality: Deserving and undeserving clients in Austrian social assistance offices. *Journal of Social Policy*, 44(1), 127–145.
- Ansell, C., Sørensen, E., & Torfing, J. (2017). Improving policy implementation through collaborative policymaking. *Policy & Politics*, 45(3), 467–486.
- Antonelli, M. A., & De Bonis, V. (2019). The efficiency of social public expenditure in European countries: A two-stage analysis. *Applied Economics*, 51(1), 47–60.
- Arts, W., & Gelissen, J. (2002). Three worlds of welfare capitalism or more? A state-of-the-art report. Journal of European Social Policy, 12(2), 137–158.

- Ascher, W. (2023). Coping with the ambiguities of poverty-alleviation programs and policies: A policy sciences approach. *Policy Sciences*, 56(2), 325–354.
- Atkinson, A. B. (2019). Measuring poverty around the world. Princeton University Press.
- Bailey, D., & Katz, J. N. (2011). Implementing panel-corrected standard errors in R: The pcse package. Journal of Statistical Software, 42(1), 1–11.
- Bakam, I., Balana, B. B., & Matthews, R. (2012). Cost-effectiveness analysis of policy instruments for greenhouse gas emission mitigation in the agricultural sector. *Journal of Environmental Management*, 112, 33–44.
- Baltagi, B. H. (2021). Econometric analysis of panel data (6th ed.). Springer.
- Besharov, D. J., & Call, D. M. (2009). Income transfers alone won't eradicate poverty. *Policy Studies Journal*, 37(4), 599–631.
- Biggeri, M., & Mauro, V. (2018). Towards a more 'sustainable' human development index: Integrating the environment and freedom. *Ecological Indicators*, 91, 220–231.
- Bonoli, G., Natili, M., & Trein, P. (2019). A federalist's dilemma: Trade-offs between social legitimacy and budget responsibility in multi-tiered welfare states. *Journal of European Social Policy*, 29(1), 56–69.
- Bozio, A., Emmerson, C., Peichl, A., & Tetlow, G. (2015). European public finances and the great recession: France, Germany, Ireland, Italy, Spain and the United Kingdom compared. *Fiscal Studies*, 36(4), 405–430.
- Brady, D., & Bostic, A. (2015). Paradoxes of social policy: Welfare transfers, relative poverty, and redistribution preferences. *American Sociological Review*, 80(2), 268–298.
- Bressers, H., & Klok, P. J. (1988). Fundamentals for a theory of policy instruments. International Journal of Social Economics, 15(3/4), 22–41.
- Cammeraat, E. (2020). The relationship between different social expenditure schemes and poverty, inequality and economic growth. *International Social Security Review*, 73(2), 101–123.
- Candel, J. J., & Biesbroek, R. (2016). Toward a processual understanding of policy integration. *Policy Sciences*, 49, 211–231.
- Cantillon, B., Marx, I., & Van Den Bosch, K. (2003). The puzzle of egalitarianism. The relationship between employment, wage inequality, social expenditure and poverty. *European Journal of Social Security*, 5(2), 108–127.
- Capano, G., & Howlett, M. (2020). The knowns and unknowns of policy instrument analysis: Policy tools and the current research agenda on policy mixes. SAGE Open. https://doi.org/10.1177/21582 44019900568
- Castles, F. G. (2004). *The future of the welfare state: Crisis myths and crisis realities*. Oxford University Press.
- Castles, F. G., & Mitchell, D. (1992). Identifying welfare state regimes: The links between politics, instruments and outcomes. *Governance*, 5(1), 1–26.
- Cejudo, G. M., & Michel, C. L. (2017). Addressing fragmented government action: Coordination, coherence, and integration. *Policy Sciences*, 50, 745–767.
- Chindarkar, N., Howlett, M., & Ramesh, M. (2017). Conceptualizing effective social policy design: Design spaces and capacity challenges. *Public Administration and Development*, 37(1), 3–14.
- Ciccia, R., & Javornik, J. (2019). Methodological challenges for comparative welfare state research: Capturing intra-country variation in cross-national analyses. *Journal of Comparative Policy Analysis: Research and Practice*, 21(1), 1–8.
- Clasen, J., & Siegel, N. A. (Eds.). (2007). Investigating welfare state change: The "dependent variable problem" in comparative analysis. Edward Elgar Publishing.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., ... & Ziblatt, D. (2019). V-dem codebook v9.
- Dasgupta, P., & Weale, M. (1992). On measuring the quality of life. World Development, 20(1), 119-131.
- De Deken, J., & Kittel, B. (2007). Social expenditure under scrutiny: The problems of using aggregate spending data for assessing welfare state dynamics. In J. Clasen & N. A. Siegel (Eds.), *Investigating welfare state change: The 'dependent variable problem' in comparative analysis* (pp. 72–104). Edward Elgar Publishing.
- Ebbinghaus, B. (2012). Comparing welfare state regimes: Are typologies an ideal or realistic strategy. Draft Paper presented at ESPAnet Conference, Edinburgh, September 6–8.
- Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276.
- Esping-Andersen, G. (1990). The three worlds of welfare capitalism. Princeton University Press.

- Fernández-i-Marín (2019): Using Policy Portfolios. Online available at http://xavier-fim.net/post/using\_ policyportfolios/.
- Fernández-i-Marín, X., Knill, C., Steinbacher, C., & Steinebach, Y. (2023). Bureaucratic quality and the gap between implementation burden and administrative capacities. *American Political Science Review*. https://doi.org/10.1017/S0003055423001090
- Fernández-i-Marín, X., Knill, C., & Steinebach, Y. (2021). Studying policy design quality in comparative perspective. American Political Science Review, 115(3), 931–947.
- Foster, J. L., Lopez-Calva, L., & Szekely, M. (2005). Measuring the distribution of human development: Methodology and an application in Mexico. *Journal of Human Development and Capabilities*, 6(1), 5–25.
- Green-Pedersen, C. (2004). The dependent variable problem within the study of welfare state retrenchment: Defining the problem and looking for solutions. *Journal of Comparative Policy Analysis*, 6(1), 3–14.
- Greve, B. (2017). How to measure social progress? Social Policy & Administration, 51(7), 1002–1022.
- Gugushvili, D., & Laenen, T. (2021). Two decades after Korpi and Palme's 'paradox of redistribution': What have we learned so far and where do we take it from here? *Journal of International and Comparative Social Policy*, 37(2), 112–127.
- Haelg, F., Potrafke, N., & Sturm, J. E. (2022). The determinants of social expenditures in OECD countries. Public Choice, 193(3–4), 233–261.
- Hagerty, M. R., Cummins, R., Ferriss, A. L., Land, K., Michalos, A. C., Peterson, M., Sharpe, A., Sirgy, J., & Vogel, J. (2001). Quality of life indexes for national policy: Review and agenda for research. *Bulletin of Sociological Methodology/bulletin De Méthodologie Sociologique*, 71(1), 58–78.
- Hanson, J. K., & Sigman, R. (2021). Leviathan's latent dimensions: Measuring state capacity for comparative political research. *The Journal of Politics*, 83(4), 1495–1510.
- Hauner, D., & Kyobe, A. (2010). Determinants of government efficiency. World Development, 38(11), 1527–1542.
- Henisz, W. J. (2000). The institutional environment for economic growth. *Economics & Politics*, 12(1), 1–31.
- Homsy, G. C., Liu, Z., & Warner, M. E. (2019). Multilevel governance: Framing the integration of top-down and bottom-up policymaking. *International Journal of Public Administration*, 42(7), 572–582.
- Hood, C. (2010). The blame game. Princeton University Press.
- Hooghe, L., Marks, G., Schakel, A. H., Osterkatz, S. C., Niedzwiecki, S., & Shair-Rosenfield, S. (2016). *Measuring regional authority: A postfunctionalist theory of governance* (Vol. I). Oxford University Press.
- Howlett, M. (2018). The criteria for effective policy design: Character and context in policy instrument choice. *Journal of Asian Public Policy*, 11(3), 245–266.
- Huber, E., & Stephens, J. D. (2001). Development and crisis of the welfare state: Parties and policies in global markets. University of Chicago Press.
- Jacques, O., & Noël, A. (2018). The case for welfare state universalism, or the lasting relevance of the paradox of redistribution. *Journal of European Social Policy*, 28(1), 70–85.
- Jacques, O., & Noël, A. (2021). Targeting within universalism. Journal of European Social Policy, 31(1), 15–29.
- Jahn, D. (2016). Changing of the guard: Trends in corporatist arrangements in 42 highly industrialized societies from 1960 to 2010. Socio-Economic Review, 14(1), 47–71.
- Jensen, C. (2011a). Less bad than its reputation: Social spending as a proxy for welfare effort in crossnational studies. Journal of Comparative Policy Analysis: Research and Practice, 13(3), 327–340.
- Jensen, C. (2011b). Catching up by transition: Globalization as a generator of convergence in social spending. *Journal of European Public Policy*, 18(1), 106–121.
- Jensen, C., Knill, C., Schulze, K., & Tosun, J. (2014). Giving less by doing more? Dynamics of social policy expansion and dismantling in 18 OECD countries. *Journal of European Public Policy*, 21(4), 528–548.
- Kautto, M. (2002). Investing in services in West European welfare states. Journal of European Social Policy, 12(1), 53–65.
- Kern, F., & Howlett, M. (2009). Implementing transition management as policy reforms: A case study of the Dutch energy sector. *Policy Sciences*, 42, 391–408.
- Kittel, B., & Winner, H. (2005). How reliable is pooled analysis in political economy? The globalization-welfare state nexus revisited. *European Journal of Political Research*, 44(2), 269–293.
- Knill, C., Schulze, K., & Tosun, J. (2012). Regulatory policy outputs and impacts: Exploring a complex relationship. *Regulation & Governance*, 6(4), 427–444.

- Knill, C., Steinbacher, C., & Steinebach, Y. (2021a). Balancing trade-offs between policy responsiveness and effectiveness: The impact of vertical policy-process integration on policy accumulation. *Public Administration Review*, 81(1), 157–160.
- Knill, C., Steinbacher, C., & Steinebach, Y. (2021b). Sustaining statehood: A comparative analysis of vertical policy-process integration in Denmark and Italy. *Public Administration*, 99(4), 758–774.
- Korpi, W., & Palme, J. (1998). The paradox of redistribution and strategies of equality: Welfare state institutions, inequality, and poverty in the Western countries. *American Sociological Review*, 63(5), 661–687.
- Levi-Faur, D. (2014). The welfare state: A regulatory perspective. Public Administration, 92(3), 599-614.
- Longford, N. T., & Nicodemo, C. (2010). The contribution of social transfers to the reduction of poverty. IZA Discussion Papers, No. 5223.
- Lundin, M. (2007). When does cooperation improve public policy implementation? *Policy Studies Journal*, 35(4), 629–652.
- Magro, E., & Wilson, J. R. (2019). Policy-mix evaluation: Governance challenges from new place-based innovation policies. *Research Policy*, 48(10), 103612.
- Malbon, E., Carey, G., & Reeders, D. (2019). Mixed accountability within new public governance: The case of a personalized welfare scheme in early implementation. *Social Policy & Administration*, 53(1), 156–169.
- Merrien, F. X. (1998). Governance and modern welfare states. *International Social Science Journal*, 50(155), 57–67.
- Metzger, N., & Shenai, V. (2022). Economic growth and human development in OECD countries: A twentyyear study of data 2000–2019. *Journal of European Economy*, 20(4), 585–631.
- OECD (2021c). Online available at: Annual Labour Force Statistics. Online available at: https://stats.oecd. org/Index.aspx?DataSetCode=ALFS\_SUMTAB.
- OECD (2021d). Family Database. Online available at: https://stats.oecd.org/Index.aspx?DataSetCode= FAMILY.
- OECD (2021a). Income Distribution Database. Online available at https://stats.oecd.org/Index.aspx?DataS etCode=IDD.
- OECD (2021b). Labour Force Statistics. Online available at https://stats.oecd.org/Index.aspx?DataS etCode=DUR\_D.
- OECD (2022). Social spending (indicator). Online available at: https://doi.org/10.1787/7497563b-en.
- Pemstein, D., Meserve, S. A., & Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, 18(4), 426–449.
- Peters, B. G. (2018). The challenge of policy coordination. Policy Design and Practice, 1(1), 1–11.
- Pierson, P. (1993). When effect becomes cause: Policy feedback and political change. World Politics, 45(4), 595–628.
- Polman, D., & Alons, G. (2021). Reap what you sow: Implementing agencies as strategic actors in policy feedback dynamics. *Policy Sciences*, 54(4), 823–848.
- Powell, M., & Barrientos, A. (2004). Welfare regimes and the welfare mix. European Journal of Political Research, 43(1), 83–105.
- Ranis, G., Stewart, F., & Samman, E. (2006). Human development: Beyond the human development index. *Journal of Human Development*, 7(3), 323–358.
- Ryan, C. M. (2001). Leadership in collaborative policymaking: An analysis of agency roles in regulatory negotiations. *Policy Sciences*, 34(3–4), 221–245.
- Sabatier, P. A., & Mazmanian, D. A. (1980). The implementation of public policy: A framework of analysis. *Policy Studies Journal*, 8(4), 538–560.
- Scharpf, F. W. (2000). Institutions in comparative policy research. Comparative Political Studies, 33(6–7), 762–790.
- Scruggs, L., Jahn, D., & Kuitto, K. (2013). Comparative welfare entitlements data set 2, Version 2013-08. http://cwed2.org/.
- Scruggs, L. (2007). Welfare state generosity across space and time. In J. Clasen & N. A. Siegel (Eds.), *Investigating welfare state change: The "dependent variable problem" in comparative analysis* (pp. 133–165). Edward Elgar Publishing.
- Sen, A. (2006). Conceptualizing and Measuring Poverty. In D. B. Grusky & R. Kanbur (Eds.), Poverty and inequality: Studies in social inequality (pp. 30–46). Stanford University Press.
- Skocpol, T. (1995). Social policy in the United States: Future possibilities in historical perspective. Princeton University Press.
- Steinbacher, C. (2023). Analytical perspectives on environmental policy integration. In H. Jörgens, C. Knill, & Y. Steinebach (Eds.), *Routledge handbook of environmental policy* (pp. 122–135). Routledge.

- Steinebach, Y. (2022). Instrument choice, implementation structures, and the effectiveness of environmental policies: A cross-national analysis. *Regulation & Governance*, 16(1), 225–242.
- Thorbecke, E. (2004). Conceptual and measurement issues in poverty analysis (No. 2004/04). WIDER Discussion Paper.
- Tosun, J., & Lang, A. (2017). Policy integration: Mapping the different concepts. *Policy Studies*, 38(6), 553–570.
- Treier, S., & Jackman, S. (2008). Democracy as a latent variable. American Journal of Political Science, 52(1), 201–217.
- Trein, P. (2020). Bossing or protecting? The integration of social regulation into the welfare state. The ANNALS of the American Academy of Political and Social Science, 691(1), 104–120.
- Trein, P., Maggetti, M., & Meyer, I. (2021). Necessary conditions for policy integration and administrative coordination reforms: An exploratory analysis. *Journal of European Public Policy*, 28(9), 1410–1431.
- Trubek, D. M., & Trubek, L. G. (2005). Hard and soft law in the construction of social Europe: The role of the open method of co-ordination. *European Law Journal*, 11(3), 343–364.
- UNDP (2023). Human Development Index and components. Online available at: https://hdr.undp.org/datacenter/documentation-and-downloads
- Valls-Fonayet, F., Belzunegui-Eraso, A., & De Andrés-Sánchez, J. (2020). Efficiency of social expenditure levels in reducing poverty risk in the EU-28. *Poverty & Public Policy*, 12(1), 43–62.
- Van Kersbergen, K. (2013). What are welfare state typologies and how are they useful, if at all? In B. Greve (Ed.), *The Routledge handbook of the welfare state* (pp. 115–123). Routledge.
- Wegrich, K. (2015). Jeffrey L. Pressman and Aaron B. Wildavsky, implementation. In M. Lodge, E. C. Page, & S. J. Balla (Eds.), *The Oxford handbook of classics in public policy and administration* (pp. 342–358). Oxford University Press.
- Wilensky, H. L. (1975). The welfare state and equality. University of California Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.