Bureaucratic Quality and the Gap between Implementation Burden and Administrative Capacities

XAVIER FERNÁNDEZ-I-MARÍN *Universitat de Barcelona, Spain*
CHRISTOPH KNILL *LMU Munich, Germany*
CHRISTINA STEINBACHER *LMU Munich, Germany*
YVES STEINEBACH *University of Oslo, Norway*

*D*emocratic governments produce more policies than they can effectively implement. Yet, this gap between the number of policies requiring implementation and the administrative capacities available to do so is not the same in all democracies but varies across countries and sectors. We argue that this variation depends on the coupling of the sectoral bureaucracies in charge of policy formulation and those in charge of policy implementation. We consider these patterns of vertical policy-process integration an important feature of bureaucratic quality. The more the policymaking level is involved in policy implementation (top-down integration) and the easier the policy-implementing level finds it to feed its concerns into policymaking (bottom-up integration), the smaller the so-called "burden-capacity gap." We demonstrate this effect through an empirical analysis in 21 OECD countries over a period of more than 40 years in the areas of social and environmental policies.

MIND THE GAP: GROWING POLICY STOCKS AND CONSTRAINED ADMINISTRATIVE CAPACITIES

Over the last decades, modern democracies have witnessed a tremendous increase in the number and complexity of public policies. Recent empirical evidence suggests that the average number of policy measures in OECD countries has grown, for instance, four times in environmental policies and doubled in social policies between the 1980s and the 2010s (Adam et al. 2019). Hurka, Haag, and Kaplaner (2022) make a similar observation. They find that the average number of articles on European Union (EU) legal documents grew more than twofold from 25 in the early 1990s to more than 75 in 2021. This trend is even more pronounced when analyzing the average word count, which skyrocketed from about 1,000 to more than 4,000 words (Hurka, Haag, and Kaplaner 2022).

More policies generally mean more implementation burdens for the authorities in charge of their execution and enforcement.[1] If policies are adopted without a parallel expansion of administrative capacities, this can directly translate into growing implementation problems (Knill, Steinebach, and Zink 2023). As Gratton et al. (2021) put it, the (over)production of public policies can shift the administration "from a Weberian to a Kafkaesque bureaucracy" if "too many and too frequent laws overload (…) the bureaucracy with too many acts to implement" (2965). The characteristic symptom of "bureaucratic overload" is that public authorities are "under-resourced relative to their responsibilities," resulting in the "failure to implement programs on a scale sufficient to meet the demand for benefits among citizens despite statutory entitlements" (Dasgupta and Kapur 2020, 1316). In a recently published opinion piece in the *New York Times*, Klein (2022) summarizes this trend as follows: US policymakers "spend too much time and energy imagining the policies that a capable government could execute and not nearly enough time imagining how to make a government capable of executing them."

Yet, the size of the gap between implementation burdens and administrative capacities seems to vary

Xavier Fernández-I-Marín [ID], Ramón y Cajal Researcher, Department of Political Science, Universitat de Barcelona, Spain, x.fernandez-i-marin@ub.edu.

Christoph Knill [ID], Full Professor, Geschwister Scholl Institute of Political Science, LMU Munich, Germany, christoph.knill@gsi.uni-muenchen.de.

Corresponding author: Christina Steinbacher [ID], Doctoral Researcher, Geschwister Scholl Institute of Political Science, LMU Munich, Germany, christina.steinbacher@gsi.uni-muenchen.de.

Yves Steinebach [ID], Associate Professor, Department of Political Science, University of Oslo, Norway, yves.steinebach@stv.uio.no.

---

[1] With "implementation burden" we explicitly capture the burdens for public implementation authorities. It refers to the additional administrative tasks to apply, monitor, and enforce the policies in place. Implementation burdens crucially differ from administrative burdens that relate to the costs people encounter "when they search for information about public services (learning costs), comply with rules and requirements (compliance costs), and experience the stresses, loss of autonomy, or stigma that come from such encounters (psychological costs)" (Herd and Moynihan 2018, 2).

considerably across countries and sectors. Levels of policy accumulation differ despite a common pattern of largely stagnating or even incrementally declining administrative capacities (Adam et al. 2019). Limberg et al. (2021, 438) substantiate this point by demonstrating that some countries can ensure that "the growth in rules [does] not outpace the expansion in administrative capacities," while others cannot (see also Fernández-i-Marín et al. 2023a).

But how can we explain this variation in the relationship between the number of policies requiring implementation and the administrative capacities available to do so? We argue that the answer to this question lies in the coupling of sectoral bureaucracies in charge of policy formulation and those in charge of policy implementation. We consider these patterns of vertical policy-process integration (VPI) an important yet relatively overlooked feature of bureaucratic quality. We expect that (1) the more the policymaking level is involved in policy implementation (*top-down integration*) and (2) the easier the policy-implementing level finds it to feed its concerns into policymaking (*bottom-up integration*), the smaller the burden-capacity gap.

To test our argument, we systematically compare the development of the gap between implementation burdens and administrative capacities across two policy areas (environmental and social policy) in 21 OECD countries over a period of more than 40 years (1976 to 2018). Our results show a very clear picture: higher levels of VPI substantially reduce the risk that the number of policies requiring implementation and the administrative capacities drift apart.

Our paper contributes to existing scholarship in three ways. Our first contribution is theoretical in nature. Building on the concept of VPI, we offer a novel theoretical approach to account for the gap between burdens emerging from sectoral policy growth and the capacities available to put these policies into effect. Although implementation research has long acknowledged the relevance of administrative capacities for implementation success (Pressman and Wildavsky 1973; Winter 2012), they have been merely assessed in absolute rather than relative terms: administrative capacities are typically classified as high or low—they are *not* assessed in reference to the implementation burden that must be handled with the capacities available. Second, we provide a conceptual approach to assess the relationship between policy accumulation and administrative capacities from a macro-perspective that goes beyond the analysis of individual instances of policy change and implementation. Third, we provide new and encompassing empirical evidence that allows for the systematic assessment of our argument across a large number of countries and a time period of over 40 years. Moreover, we demonstrate our argument by studying the impact of VPI on the burden-capacity gap in two highly diverse sectors of social and environmental policies.

The remainder of this paper is organized as follows. In the subsequent section, we present our theoretical argument. This will be followed by an introduction of our research design before we turn to the presentation and critical discussion of our empirical findings. The final section concludes and highlights the implications of our results for future research and practice. Supplementary material provides complementary information on our quantitative analysis and the qualitative substantiation of our arguments.

EXPLAINING THE GAP: THE IMPACT OF VPI

If policymaking was merely a matter of politics, the outlook for the development of the burden-capacity gap would be rather gloomy. Power-seeking politicians have strong incentives to demonstrate their responsiveness to societal demands by constantly proposing new policies (Gratton et al. 2021).[2] The same logic, however, does not apply to the expansion of administrative capacities that are needed for properly implementing these new policies. As political responsibilities for implementation success are often unclear, electoral incentives for politicians to invest in administrative capacities for policy implementation are generally weaker than those for adopting new policies (Dasgupta and Kapur 2020). Although such capacity expansion improves implementation effectiveness, attributing such improvements to the actions of particular political actors is in many instances more difficult for voters (Hinterleitner 2020). While political actors have strong incentives to engage in policy production, their incentives to engage in costly improvements of administrative capacities are typically much weaker. Deviations from this pattern are largely confined to specific constellations where implementation failures have immediate individual consequences for voters, for example, in the case of policies related to service delivery (Healy and Malhotra 2013).

From a mere political logic, we should hence expect a toxic combination of strong policy growth and stagnating or even declining administrative capacities. Although the urgency of this problem varies across countries and sectors, the nature of the challenge essentially remains the same: if we assume that policymaking is *exclusively* driven by politics, we would expect an ever-growing burden-capacity gap, with more and more policies *undermining* rather than strengthening overall policy effectiveness in the long run.

Yet, politics is not the only factor at play. Politics might be counterbalanced by bureaucratic processes, as politicians strongly depend on bureaucracies' information and expertise when developing new policy proposals (Nicholson-Crotty and Miller 2012). Bureaucracies' potential to effectively attenuate the responsiveness-driven logic of politics by concerns of policy effectiveness cannot be taken for granted,

---

[2] There are countless political opportunities for policy production. While voters may hold policymakers accountable for new policies causing large budgetary deficits, there are numerous strategies available to address societal needs through additional policies that have minimal budgetary implications, such as regulatory policies.

however. Challenges emerge from the high degree of functional and structural differentiation characterizing modern bureaucracies. Governmental functions are divided not only across policy sectors but also across the stages of the policy cycle (e.g., policy formulation and implementation). Moreover, functions are assigned to different levels of government (Hooghe and Marks 2003). This structural decoupling implies that policy-formulating bureaucracies may shift the implementation costs of the policies they produce to other administrative bodies or levels of government. Similarly, implementation bodies may find it difficult to communicate their policy experiences and needs from the lower administrative echelons up to the policymaking level.

To understand the balance of accumulating implementation burdens and administrative capacities, greater attention needs to be paid to the coupling of administrative bodies responsible for policy formulation and those in charge of policy implementation. While the tasks of policy formulation typically lie with bureaucracies at the ministerial level, "the polity of implementation" (Sager and Gofen 2022) follows a more diverse pattern of national and sectoral variation involving bureaucracies at very different levels of government.[3]

We refer to these bureaucratic coupling arrangements as VPI. Our conception of VPI departs from a multilevel governance perspective that focuses on the interaction processes both between and within different layers of government (Peters and Pierre 2004). We conceive VPI as a specific feature of bureaucratic quality. Bureaucratic quality can generally be defined as the "degree to which government policies are constructed and implemented accurately, swiftly, competently, and impartially" (Andersen 2018, 247; see also Hanson and Sigman 2021). To assess bureaucratic quality, the focus has been mainly on other structural features, in particular the degree of bureaucratic professionalism, impartiality, and autonomy (Fukuyama 2013). While these aspects affect macro-level outcomes like socioeconomic development (Rauch and Evans 2000), corruption prevention (Dahlström, Lapuente, and Teorell 2012), and popular satisfaction (Dahlberg and Holmberg 2014), their potential impact on aggregate policy dynamics, such as the burden-capacity gap, is much less obvious.

VPI affects the burden-capacity gap in two ways. First, VPI defines the leeway that policy-producing bureaucracies have for unloading implementation costs to subordinate bodies (*top-down integration*). Second, the systematic integration of implementers' experience into policy formulation (*bottom-up integration*) informs policymakers regarding what works in practice and what is needed to further improve policy design and implementation.

[3] In some countries, environmental policies, for instance, are implemented by central environmental protection agencies, whereas in others, vast shares of implementation tasks are performed by subnational entities.

Top-Down Integration

As policies are rarely self-implementing, they usually come with some burden of implementation: Services must be provided, policies enforced, and compliance monitored. In short, implementation is costly, and these costs typically remain in the realm of public budgets, even when the government decides to delegate implementation tasks to private sector bodies (Ansell and Gash 2008).

Yet, the tasks of policy formulation and implementation are often located at different levels and places of government. As shown in the literature on US federalism, central governments tend to impose additional financial and administrative burdens on subnational governments via unfunded mandates (Moffitt et al. 2021). There is hence a considerable potential that the costs and benefits of new policies are decoupled. On the one hand, the policy formulation level may benefit from demonstrating responsiveness to societal demands. On the other hand, the burden to apply and enforce these new measures accumulates at the implementation level. In such setups, there are relatively few barriers to the constant overproduction of policies. The top-down dimension of VPI takes this as a starting point (Knill, Steinbacher, and Steinebach 2021). It determines the leeway that policy producers have to unload the cost of policy implementation onto other administrative levels or bodies. We argue that this leeway varies with the factors: (1) accountability, (2) responsibility for administrative resources, and (3) organizational setup.

Accountability captures the extent to which policymakers can be held politically responsible for policy implementation. This depends not only on the formal arrangements that determine the extent to which policymakers must carry the burden of exercising legal and administrative oversight over implementation bodies (Hill and Hupe 2009) but also on informal opportunities to shift blame for implementation failures to other actors (Hinterleitner 2020). The responsibility for administrative resources, by contrast, refers to the allocation of costs associated with the policy and its administration. Direct policy costs, the costs of child benefits, for example, must be accounted for. Moreover, implementers must be trained, employed, and equipped. Finally, the organizational setup captures a third type of implementation cost that relates to the responsibility for setting up and designing implementation structures (Winter 2012).

The distribution of these different implementation costs between policy formulation and implementation levels can significantly differ across countries and sectors. Higher degrees of top-down integration imply that the policy-formulating bureaucracies must also bear the implementation costs of the policies they produce. This way, top-down integration can be expected to reduce the scope of the burden-capacity gap in two ways. Higher levels of top-down integration may (1) curb the growth of implementation burdens on the ground by producing fewer policies and (2) increase the chances that those policies

produced are backed with administrative capacities for implementation.

Bottom-Up Integration

There are numerous reasons why policies might fail to achieve their intended results. Policies might suffer from design flaws such as overly ambiguous objectives (Pressman and Wildavsky 1973), faulty assessments of the nature of the problem (Linder and Peters 1984), or incorrect assumptions about means-ends relationships (Schneider and Ingram 1997). Moreover, governments might not possess the capacities required by a given policy instrument (Howlett and Ramesh 2016). While all these design flaws are essentially created during the policy formulation process, they are often identified only at the implementation stage, that is, by implementing bodies whose daily work reveals such discrepancies between policies' intentions and actual conditions, as well as between the capacities needed and those available.

Against this background, the information flow from the bottom (policy implementation) to the top (policy formulation) emerges as an important condition affecting the design of effective policies. Yet, managing information flows in highly differentiated politico-administrative systems is a demanding undertaking, and its success cannot be taken for granted. Although "street-level policy entrepreneurship" thus constitutes an important part of policymaking (Cohen 2021), the actual realization of this potential depends on various institutional factors. These are: (1) articulation, (2) consultation, and (3) evaluation.

First, effective involvement in policy formulation presumes articulation. Articulation captures the implementation level's ability to develop clear and coherent positions when assessing different policy design options. The bottom-up development of design options is often hindered by the fact that implementation bodies must juggle competing and sometimes contradictory demands (Cohen and Frisch-Aviram 2021). Articulation hence requires organizational integration across different implementation bodies, such as umbrella organizations or other platforms of local or regional authorities representing or actively advocating the interests of their member organizations. Yet, effective participation in policy formulation requires not only the formulation of common positions but also that these positions are actually taken into account. This aspect is reflected in the consultation indicator, which captures the presence and extent of institutional arrangements that allow implementation bodies to present their concerns and positions to the bureaucracies responsible for designing new policy proposals. The latter are of particular importance as implementing bodies typically have few direct formal channels of communication or close informal relationships with high-level bureaucrats and politicians (Cohen 2021). Third, opportunities to learn from the implementers' experience increase with the existence and usage of systematic ex-ante and ex-post evaluations. Bottom-up integration in this context does not necessarily require that evaluations are merely policy-based and isolated from political controversies. Regardless of the concrete setup and the extent of politicization, the existence of evaluation practices increases the chance that the implementation level gets a voice within policy formulation processes (Stockmann, Meyer, and Taube 2020).

Taken together, higher degrees of bottom-up integration should ensure two things. First, implementation bodies can signal to policy formulators the required amount and type of resources needed to make policies work in practice. Second, the inclusion of policy implementers' experience into policy formulation might help to improve policy design and hence overall policy effectiveness. This, in turn, reduces the need for new policies to correct the errors and deficiencies of already existing measures. Consequently, the growth of public policies and associated implementation burdens should be dampened.

VPI and the Burden-Capacity Gap

In the previous sections, we discussed two constituent dimensions of VPI and how they affect the gap between accumulating implementation burdens and available administrative capacities. Careful readers might argue that the two dimensions under scrutiny do not affect the burden-capacity gap in general but rather exert an individual influence on either the development of the implementation burdens (mainly by top-down integration) or on available administrative capacities (mainly by bottom-up integration). In practice, however, it is hardly possible to disentangle the distinct effects of VPI dimensions on implementation burden and administrative capacities. Bottom-up integration helps implementation bodies to effectively communicate to policymakers that they lack administrative capacities. Yet, this might lead policymaking bodies to either provide more resources or to not put additional tasks on their shoulders. In a recent example in Germany, local authorities published a public letter stating that their "load limit has been exceeded" (Gemeindetag Baden-Württemberg 2022, 4) and that, in consequence, "new standards, legal rights, and benefits can no longer be implemented." They conclude by calling the upper levels of government to engage in an "honest and serious discussion" on what the state can actually provide. Likewise, top-down integration might not only lead to the production of fewer policies overall (as policymakers must carry implementation costs) but also to a stronger coupling of the development of implementation burdens and capacities, as policymakers are fully accountable for the implementation success of the policies they adopted. In essence, this implies that understanding the burden-capacity gap requires a holistic approach that considers the combined effects of multiple factors on both burden *and* capacity, rather than isolating them as separate entities. The VPI concept purportedly provides such a perspective by explicitly focusing on the gap phenomenon

FIGURE 1. Central Considerations of VPI and Its Effect on the Burden-Capacity Gap
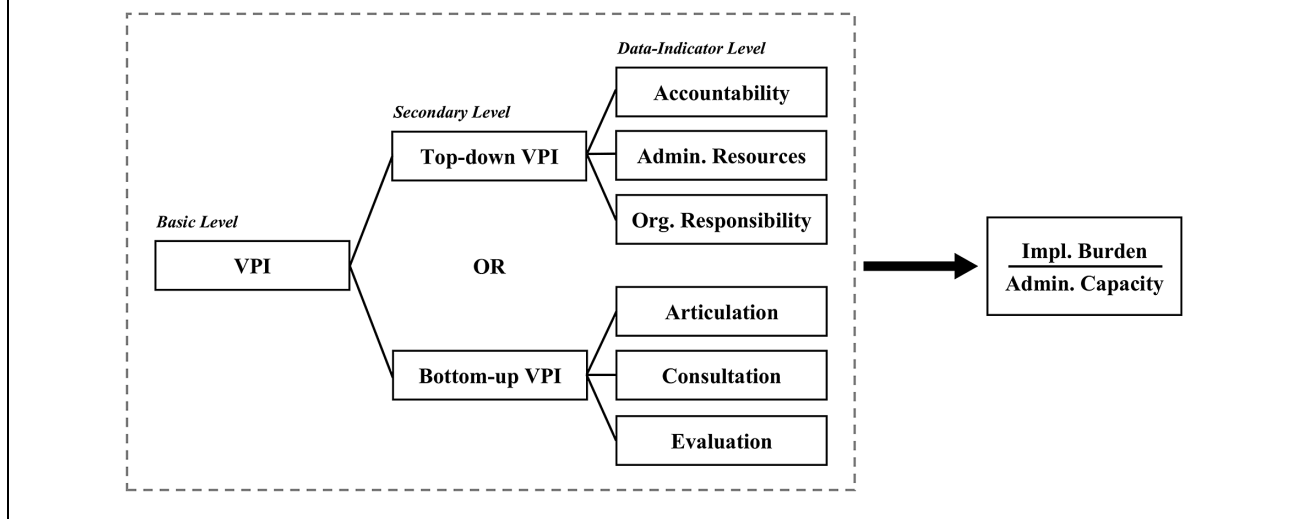
and recognizing the complex relationship between its determinants.[4]

As we have no prior knowledge about the exact causal effect of top-down and bottom-up integration on the burden-capacity gap, we assume substitutability among our two dimensions as the underlying logic of concept formation. This means that the two dimensions are not linked by the operator "and" (necessity) but by the operator "or" (substitutability) (Goertz 2006, chap. 2). This implies that we expect top-down and bottom-up integration to commonly constitute and represent the degree of VPI and to jointly affect the burden-capacity gap. Therefore, our central expectation is that higher levels of sectoral VPI reduce the gap between implementation burdens emerging from policy accumulation and available administrative capacities.

Our central conceptual and theoretical considerations on VPI and its effect on the burden-capacity gap are summarized in Figure 1. Following Goertz's (2006) logic for concept formation, our model is structured in three hierarchical levels: at the *basic level*, we have VPI serving as the core concept; the *secondary level* expands on this by introducing the aspects of top-down and bottom-up VPI; and lastly, the *data-indicator level*, detailed further in the section on measuring VPI, comprises articulation, consultation, evaluation, accountability, administrative resources, and organizational responsibility. In this framework, we posit an inverse correlation between VPI and the burden-capacity gap, suggesting that increasing VPI could potentially reduce this gap.

RESEARCH DESIGN

Our country sample comprises 21 OECD countries (see Figure 3). While all the countries under analysis constitute advanced industrialized democracies, the institutional characteristics of their political systems vary substantially. This allows us to study the impact of VPI across a set of (otherwise) very different institutional conditions and, as such, to increase the generalizability of our findings (Gerring 2008). The investigation period extends from 1976 to 2018.

We focus on the policy areas of environmental and social policies. Within each policy area, we study three policy fields: in environmental policy—air and climate, water, and nature protection policies; in social policy—pensions, unemployment, and child benefits. Studying both environmental and social policies allows us to test our argument across: (1) different policy types (regulatory versus redistributive policies); (2) fields with different degrees of maturity (young versus old fields); (3) quite different requirements for implementation (authorization, inspection, and planning versus public service provision); and, relatedly, (4) differences in political incentives for engaging in capacity expansions in view of voters' affectedness in case of implementation failures (diffuse matters of environmental quality versus problems in individual service delivery). If our argument holds across this wide range of contextual conditions, it should also apply to other policy areas and constitute a general feature of bureaucratic quality.

Sizing the Burden-Capacity Gap

Our dependent variable is the gap between the size of the implementation burden as indicated by the level of policy accumulation, on the one hand, and administrative capacities, on the other. To assess the "size" of this gap, we calculate the ratio between the level of policy accumulation and the administrative capacities available. The higher the value of this

---

[4] In section "Empirical Findings and Discussion" and Figure A17 of the Supplementary material, we provide additional empirical evidence, comparing the impact of VPI dimensions on the gap and its constituent components.

quotient, the larger the burden-capacity gap. To make this measurement work in practice, we need a sound conceptualization of both the level of policy accumulation and the administrative capacities available.

Implementation Burden

We approach the development of the size of the policy-induced implementation burden by tracing changes in the size of sectoral policy portfolios over time (see Adam et al. 2019). Policy portfolios are typically composed of two dimensions: policy targets and policy instruments. Policy targets are all issues addressed by governments. Policy instruments are the means governments can use to meet these targets. The differentiation between policy targets and instruments leaves us with a two-dimensional portfolio space. Based on this portfolio space, we can calculate a standardized measure of the sectoral portfolio size that can range from 0 (no policy instrument for any target) to 1 (all policy instruments for all targets).

For illustrative purposes, let us imagine a hypothetical world in which there are only two environmental issues: carbon dioxide (CO_2) emissions from cars and CO_2 emissions from industrial plants. In this simplified world, the government is also restricted to only two means to reduce CO_2 emissions: green taxation and an obligatory emission limit. When this hypothetical government does all it can to fight climate change, it addresses both targets (CO_2 emissions from cars and CO_2 emissions from industrial plants), each time using its entire policy toolkit, namely a tax and an obligatory emission limit. In this case, we would record a policy portfolio size of 1. If the government addresses all possible policy targets with only one of the two available instruments, we would record a policy portfolio size of 0.5, given that only half the possible combinations are covered. The same scenario would apply if our hypothetical government addressed only one of the two potential targets but used all the instruments in its toolkit. Unfortunately, in the real world, policymakers must deal with multiple policy issues. Luckily, however, they also possess more policy instruments to address those issues.

We measure policy-induced implementation burdens with reference to a benchmark of a (conceptually) maximum number of policy targets and instruments for each policy field under study. We predefined the relevant policy targets and instruments based on other scholarly contributions and the information provided by international organizations, public authorities, and non-governmental organizations. In the area of environmental policy, we identify 50 policy targets across the three subfields of clean air, water, and nature conservation policy. Policy targets are mostly pollutants like ozone, carbon dioxide, or sulfur dioxide in the air but also comprise other substances like lead content in gasoline, sulfur content in diesel, and nitrates and phosphates in continental surface water, as well as environmental objects like native forests, endangered plants, and endangered species. We distinguish 12 types of policy instruments (plus one residual category) that

range from hierarchical forms of governing, such as obligatory policy standards and technological prescriptions, to economic (dis-)incentives through taxes, subsidies, and other forms of market intervention. For social policy, we distinguish a total of 27 policy targets spread across the three subfields of unemployment, retirement, and children and seven instrument types, including, for instance, permanent allowances, one-time bonuses, tax exemptions, retention periods, and contributions. A list of all policy targets and instruments is provided in section A of the Supplementary material.

To illustrate our approach more explicitly, Figure 2 presents the Portuguese environmental policy portfolio at two points in time (1976 and 2018). It shows how the policy portfolio in the area of environmental policy has increased from 2 percent of the total space occupied in 1976 to 33 percent in 2018. The boxes marked in gray represent new environmental policy targets and instruments added to the portfolio.

The data on policy targets and instruments in place has been collected within the CONSENSUS[5] and the ACCUPOL[6] projects.[7] Changes in policy targets and policy instruments were assessed by scrutinizing all relevant national legislation that has been adopted throughout the observation period. The assessment hence relies on a comprehensive data collection encompassing all relevant national legal documents—laws, decrees, and regulations—in the specific policy areas under review. The legislation was collected from national legal repositories. Analysis of legislative content and the coding was carried out and cross-validated by the project teams. Additional checks of data reliability were performed based on legal commentaries and secondary literature. A coding manual (see section A of the Supplementary material) helped to extract the relevant information from the legal documents. Our measurement of policy-induced implementation burdens is therefore derived from the number of policy instruments and policy targets. An illustrative overview of the temporal evolution of the 21 national sectoral policy portfolios proxying implementation burden over time can be found in Figure A2 of the Supplementary material.

Administrative Capacities

Administrative capacities can imply very different things in practice (Lodge and Wegrich 2014). Some policies require public authorities to deliver policies themselves (delivery capacity). Other policies, in turn, require public authorities to regulate, monitor, and enforce policy delivery by non-state actors (regulatory capacity). Some policies largely depend on the

---

[5] Full project title: "Confronting Social and Environmental Sustainability with Economic Pressure: Balancing Trade-offs by Policy Dismantling or Expansion?"
[6] Full project title: "Unlimited Growth? A Comparative Analysis of Causes and Consequences of Policy Accumulation."
[7] The original dataset on environmental and social policies is published on https://publicpolicy-knill.org.

FIGURE 2. Exemplary Policy Portfolio
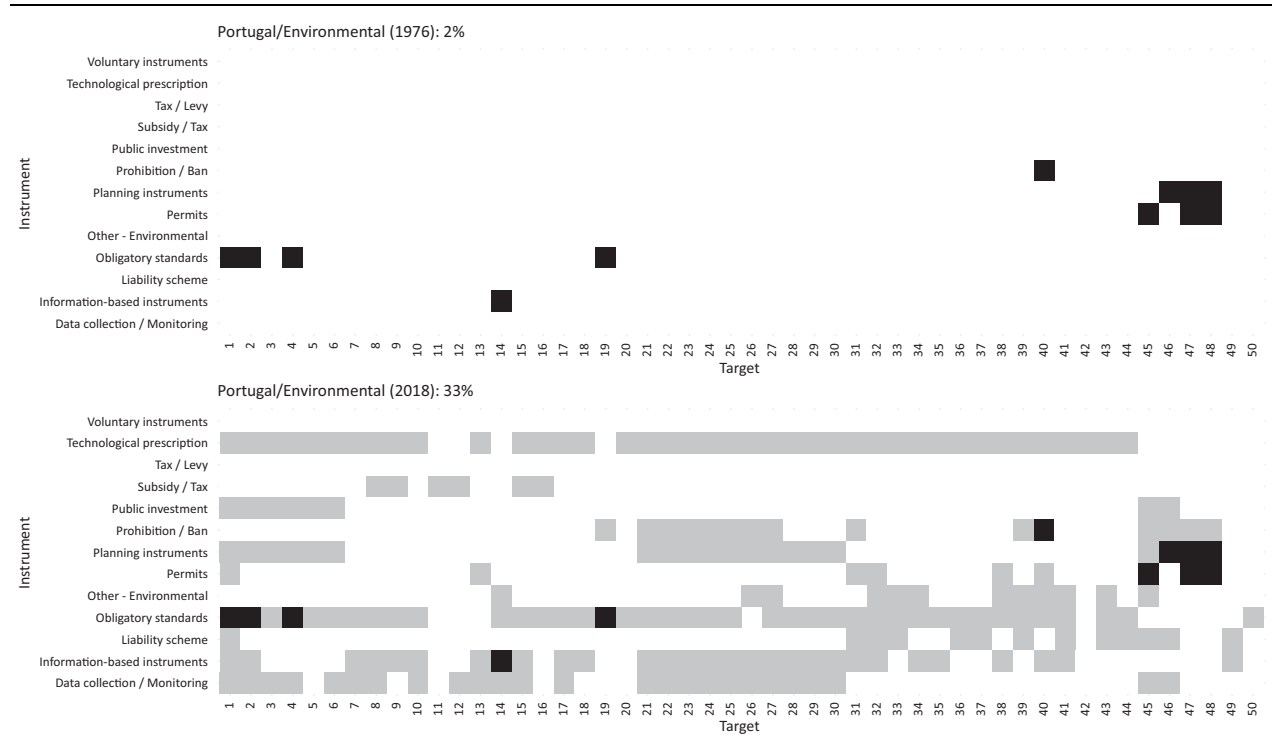
Note: The figure illustrates the policy portfolio approach. The boxes marked in gray represent the new environmental target-instrument combinations added to the portfolio between 1976 and 2018.

administrations' ability to coordinate and mediate between various actors and authorities. Moreover, policies might require the ability to evaluate the state's courses of action within highly complex environments (analytical capacity). Lastly, the success of some policies—in particular, taxes, charges, and fees—depends on the administrations' capacity to extract resources from society and manage the respective revenue streams (extractive capacity) (Bäck and Hadenius 2008).

By taking an aggregate view of changes in sectoral policy portfolios and the corresponding implementation tasks, we also require a rather broad-based measure that encompasses the different dimensions of administrative capacities. Unfortunately, most of the available measures suffer from shortcomings or capture only some dimensions of administrative capacities (Kaufmann, Kraay, and Mastruzzi 2011). We decided to construct our own measure of administrative capacities using a combination of different data sources (for a similar approach, see Fernández-i-Marín et al. 2023a; 2023b; Hanson and Sigman 2021). In essence, these provide information on the quality of national public administration provided by the V-Dem 10 dataset (Coppedge et al. 2019) and the World Bank (2023). Moreover, we include the Weberianness index by Rauch and Evans (2000) and the index of information capacity as generated by Brambor et al. (2020). In sum, our approach relies on the assumption that

administrative capacity is a multifaceted phenomenon and that only through the combination of different data sources we can come close to the actual administrative capacities available. Table 1 provides a summary of different data sources and how they are transformed for the final administrative capacity score.

In addition to these general measures of administrative capacities, we adjust our measure for sectoral differences and peculiarities. This indicates how much emphasis the government puts on the implementation of environmental or social matters. For environmental policy, we use the environmental institutional capacity index provided by Jahn (2016a). This index combines information on different dimensions of institutional capacity, such as the existence of specialized governmental institutions (e.g., environmental ministries or agencies), fundamental legal infrastructure (e.g., environmental information acts), and institutions for sustainable development (e.g., sustainability councils). For social policy, we refer to the government's spending on public employment services and administration in the area of labor market policy, including placement, counseling, and vocational guidance. These data can be readily obtained from the Social Expenditure Database of OECD (2020). While this database provides information on various kinds of social spending (e.g., old age, family, or health), it explicitly reports administrative expenses of labor market policy implementation only.

TABLE 1. Variables and Data Sources Used to Construct the Administrative Capacity Score

| Variable	Description	Dimensions	Source	Transformation	Coverage
Rigorous and impartial public administration	Captures the extent to which public officials generally abide by the law and treat like cases alike, or, conversely, the extent to which public administration is characterized by arbitrariness and bias	Delivery and regulatory capacity	Coppedge et al. (2019)	None	Broad
State authority over territory	Captures the extent of recognition of the preeminent authority of the state over its territory. It assesses the areas over which it is hegemonic, for example, where it is recognized as the preeminent authority and can, if needed, assert its control over political forces that reject its authority	Delivery and regulatory capacity	Coppedge et al. (2019)	Log (100+1-value)	Broad
Bureaucratic remuneration	Captures the extent to which state administrators are salaried employees. A state administrator is anyone who works for the state administration. A salaried employee is someone who is employed on a contract and paid a regular fee directly out of the state coffers	Delivery and regulatory capacity	Coppedge et al. (2019)	None	Broad
Criteria for appointment decisions in the state administration	Captures the extent to which appointment decisions in the state administration are based on personal and political connections, as opposed to skills and merit. Appointment decisions include hiring, firing, and promotion in the state administration	Delivery and regulatory capacity	Coppedge et al. (2019)	None	Broad
Tax revenue (% of GDP)	Measures the overall extractive capacities of the state	Extractive capacity	World Bank (2023)	Log (value-min)	Broad
Taxes on income, profits, and capital gains (% of revenue)	Captures taxes that are administratively more complex, thus requiring higher levels of record keeping and transparency and a more sophisticated bureaucratic apparatus than other revenue sources	Extractive capacity	World Bank (2023)	Log (value-min)	Broad
Taxes on international trade (% of revenue)	Captures taxes that are administratively much easier to collect and, like rents from mineral resources, do not require significant enforcement capacity. Taxes on international trade (% of revenue) are thus weighted negatively	Extractive capacity	World Bank (2023)	Log (value-min)	Broad
Statistical capacity score	Assesses the capacity of a country's statistical system. It is based on a diagnostic framework assessing methodology, data sources, periodicity, and timeliness	Analytical capacity	World Bank (2023)	None	Broad
Informational capacity	Assesses the capacity of a country to collect and process information. This covers the regular implementation of a reliable census, the regular release of statistical yearbooks, the introduction of civil and population registers, and the establishment of a government agency tasked with processing statistical information	Analytical capacity	Brambor et al. (2020)	None	Broad

(Continued)

TABLE 1. (*Continued*)

Variable	Description	Dimensions	Source	Transformation	Coverage
Weberianness	Measures the degree to which agencies employ meritocratic recruitment and offer predictable, rewarding, long-term careers	Delivery and regulatory capacity	Rauch and Evans (2000)	None	Broad
Environmental institutionalization	Assesses the degree of environmental institutionalization in a country. This combines information on different dimensions of institutional capacity, such as the existence of specialized governmental institutions (e.g., environmental ministries or agencies), fundamental legal infrastructure (e.g., environmental information acts), and institutions for sustainable development (e.g., sustainability councils)	Sectoral capacity in environmental matters	Jahn (2016a)	None	Sector-specific
Administrative spending on labor market policy per population	Captures the governmental spending on public employment services and administration in labor market policy, including placement, counseling, and vocational guidance	Sectoral capacity in social matters	OECD (2020)	Log (value+1)	Sector-specific

We performed a Bayesian latent-variable model to combine different sources into two related scores of administrative capacities. Both scores contain broad, general variables and one additional sector-specific factor. The first score contains the specific variable for the environmental sector (environmental institutionalization), while the second score contains the social sector-specific one (administrative spending on labor market policy). The model produces posterior estimates of the latent construct (a state's administrative capacity) by averaging information from each source in a similar way as a factor analysis.

Bayesian latent variable models provide several advantages when dealing with datasets of varying quality. They effectively handle missing data issues, account for measurement errors through explicitly modeling relationships between observed and latent variables, and enable the estimation of credible intervals or posterior distributions to capture parameter estimate uncertainty, even in the presence of varied data quality. Furthermore, these models capture dependencies, correlations, and hierarchical structures within the data, making them well-suited for handling datasets with complex structures and varying levels of reliability (Reuning, Kenwick, and Fariss 2019).

We apply a geometrical loss function and use the parameters' posterior means as our point estimates for the final scores. Section C of the Supplementary material (Figure A3) presents the temporal evolution of the administrative capacity scores for each country and sector. Here, we also report several indicators of reliability, including correlations between the variables and the resulting score and the discriminations of the original variables. The consistency of the resulting score of administrative capacity is supported by a congeneric reliability (omega) value of 0.78 in the environmental sector and 0.79 for the social sector.

A challenge for our analysis is that the ratio between the level of policy accumulation and available administrative capacities can be meaningfully interpreted only with reference to an empirical benchmark or anchor. We do so by using the mean value along both dimensions as a reference value. More precisely, our sectoral measures of portfolio size and administrative capacity are centered at 0 (mean value) and scaled at one standard deviation. By also testing our models with other ways of calculations, we ensured that the transformation did not affect our key results (see section E.3, Figure A11 of the Supplementary material). As a consequence of this transformation, most of our observations range between −2 and +2. A value of 0 indicates that the numerator and the denominator are at their global mean or jointly move in one direction (smaller portfolios and lesser capacities or greater portfolios with bigger capacities). Finally, we additionally validate our concept and measure of the burden-capacity gap by showing that the size of the gap between sectoral implementation burdens and available capacities is indeed negatively associated with sector-specific policy performance: A widening burden-capacity gap generally decreases the effectiveness of public policies as reported in section F of the Supplementary material.

FIGURE 3. The Burden-Capacity Gap for 21 OECD Countries (1976–2018)
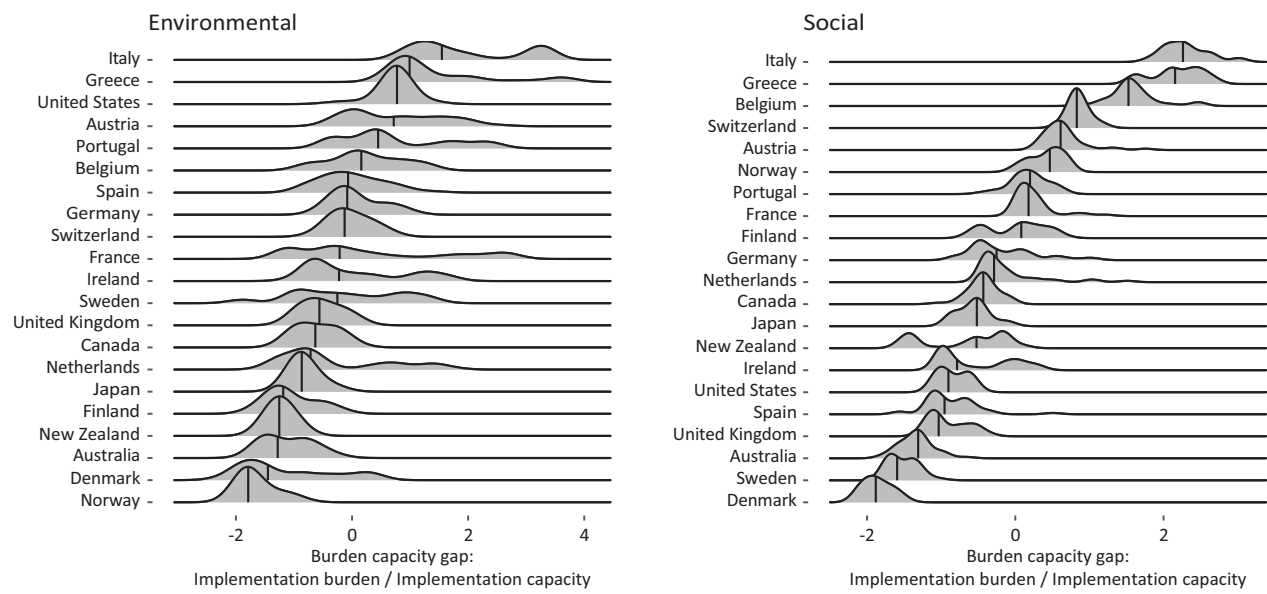
Note: The figure represents the distribution of the burden-capacity gap for each country. The vertical lines show the median.

At a value of around 1 of the burden-capacity gap, additional policy measures prove to be largely ineffective.

Figure 3 presents the burden-capacity gap for all countries under analysis. In environmental policy, the country with the largest average gap is Italy (1.97), while the country with the smallest average gap is Norway (−1.65). In social policy, Italy has again the largest average gap (2.30), while Denmark has the smallest (−1.86).

Measuring VPI

As discussed above, we expect the burden-capacity gap to be primarily affected by the degree of VPI. VPI, in turn, consists of two dimensions: top-down integration and bottom-up integration. As discussed above, both dimensions have the potential to affect either the number of policies produced or the administrative capacities available. Given that both dimensions can theoretically have the same effect, we simply aggregate the two dimensions, that is, logical "or" (Goertz 2006). The logic of substitutability also applies to our indicators manifesting either the degree of top-down integration (accountability, administrative resources, and organization) or bottom-up integration (articulation, consultation, and evaluation). In practical terms, this implies that, for instance, a lack of policy evaluation efforts can be compensated by well-organized implementation authorities who are able to articulate and advocate a common position. At the same time, positive (e.g., *more* formal accountability) and negative

(e.g., *less* consultation opportunities) changes can offset each other.

Overall, we have six different indicators capturing the manifestations of VPI's influence on the distribution of implementation costs (accountability, administrative resources, and organization) and policy design (articulation, consultation, and evaluation). Depending on the exact institutional setup and the actual reliance on these integrative channels, each indicator can take on the values of 0 (low), 0.5 (medium), or 1 (high). In sum, our VPI index thus ranges from a theoretical minimum value of 0, representing the complete absence of vertical integration, to a maximum value of 6, indicating a full integration. Table 2 presents indicators and empirical examples for the top-down and Table 3 for the bottom-up VPI dimensions.

The task at hand is to strike a balance between abstractness and specificity when measuring a latent concept, such as VPI. On the one hand, it requires predefining a set of items (see indicators below) that can generate reasonably accurate estimates for each dimension, while encompassing the overall concept. These items should capture abstract dimensions related to VPI, refer to observable states in the real world, and be as clear as possible when applied to specific cases. On the other hand, coding particular cases necessarily and inevitably involves expert judgment, no matter how clearly an indicator is formulated (Hooghe, Marks, and Schakel 2008). To strike a balance between these different requirements, our coding methodology for national VPI patterns is underpinned by a meticulous analysis and comparison of secondary literature, along with official legal, statutory, and organizational documents of

---

TABLE 2. Indicators and Examples of Top-Down VPI

With regard to (1) accountability, we coded:
'0': The policy formulation level has no formal accountability for policy implementation and has ample opportunities for political blame shifting.
> Example: *Environmental policymaking in Greece and Italy (until 2008) exemplifies weak formal accountability at the policy formulation level. Fragmented and overlapping jurisdictions highly complicate responsibility attribution.*

'0.5': The policy formulation level is partially formally accountable for policy implementation, with mechanisms of legal and administrative supervision being only partly developed.
> Example: *In the United Kingdom, with the devolution of environmental competencies as of 1998, Whitehall retained legal oversight over national framework legislation, but no competence for technical-administrative supervision.*

'1': The policy formulation level is formally accountable for policy implementation, can rely on comprehensive mechanisms of legal and administrative supervision, and has limited opportunities for political blame shifting.
> Example: *In France, "prefects" personify central accountability. Another example is Ireland's social policy administration, which became subjected to rigorous central oversight in response to the financial crisis.*

With regard to (2) administrative resources, we coded:
'0': The policy formulation level does not have to carry the policy and administrative costs of policy implementation.
> Example: *In Germany, the authorities at the Länder level are the main implementation bodies for environmental policies. Their implementation costs, however, are not covered by the federal level. Similarly, in France, the central government had no resource responsibility for implementing social security schemes prior to the Generalized Social Contribution Tax and the Juppé Plan.*

'0.5': Policy and administrative costs of policy implementation are shared between the policy formulation and the implementation level.
> Example: *In the Danish social and the British environmental sector (since 1998), policy implementation has been co-financed whereby local governments compensate those costs that exceed the fixed recoverable expenses.*

'1': The policy formulation level has to carry the full policy and administrative costs of policy implementation.
> Example: *In US social policy, the Unfunded Mandates Reform Act 1995 restored federal responsibility for the provision of all costs related to the execution of centrally adopted policies.*

With regard to (3) organization, we coded:
'0': The implementation level is fully in charge of the organization of implementation (competence allocation, setting up of organizational structures).
> Example: *In Finnish social and Dutch environmental policy (1986-1994), there are no ministerial authorities but autonomous central agencies responsible for most implementation tasks (assigned value "0").*

'0.5': Responsibility for the organization of implementation varies between policy formulation and implementation levels or is shared by both.
> Example: *The German Federal Environment Agency collaborates with state-level organizations to support implementation by state or local authorities.*

'1': The policy formulation level is fully in charge of the organization of implementation (competence allocation, setting up of organizational structures)
> Example: *In Irish and Portuguese social policy, central ministerial agencies carry out key implementation tasks (social service delivery) themselves.*

---

varying origins, such as parliamentary, governmental, or administrative. We further bolstered our methodology through consultations with country-specific experts, namely, scholars specializing in political and administrative sciences. The project members, the authors of this study, undertook the management and supervision of a team of five to six country specialists, ensuring a high degree of consistency and validity in their coding practices. Each indicator evaluation was cross-verified twice: once against a case possessing the same indicator value, and once against a case with a different indicator value. This dual verification process ensured uniformity in assessing cases with identical setups while guaranteeing distinct coding for cases with different circumstances. In conclusion, *every* data point included in our analysis represents a mini case study, revealing insights into the unique VPI patterns within a country or sector at a given point in time.[8]

Figure 4 presents the VPI score per sector and country over time. The figure reveals that the VPI varies (1) across countries, (2) over time, and (3) even between policy sectors within the same countries. Moreover, it reveals that (4) the VPI goes beyond the mere structural features of the state. While, for

---

[8] The original dataset on VPI is published on https://publicpolicy-knill.org.

TABLE 3.   Indicators and Examples of Bottom-Up VPI

With regard to (1) articulation, we coded:
'0': Implementation bodies are weakly organized and fragmented.
>   Example: *In the Australian environmental administration, there were no coordinating platforms until the creation of the Council of Australian Governments and the Inter-Governmental Agreement on the Environment in 1992.*

'0.5': Implementation bodies are well organized but often separated by policy issues and not aligned.
>   Example: *In the United States, there are several well-organized representative umbrella organizations, such as the National League of Cities, but these organizations are confronted with budgetary constraints or fierce internal polarization.*

'1': Implementation bodies are well organized and articulate coordinated positions.
>   Example: *Local Government Denmark KL takes an active part in financial negotiations with the central government and serves as an initiator for policy formulation and cooperation.*

With regard to (2) consultation, we coded:
'0': Implementation bodies are not involved in the policy formulation process.
>   Example: *In the early years of Dutch environmental policy, consultations with implementation bodies were side-lined by a strong focus on centralized command-and-control approaches.*

'0.5': Implementation bodies are sometimes involved in the policy formulation, but this varies from case to case.
>   Example: *The first National Environmental Policy Plan in the Netherlands (1989) created opportunities for participation in policymaking at the local level.*

'1': Implementation bodies are typically consulted in the policy formulation process.
>   Example: *In Finland, it is formally prescribed to consult with implementation bodies and affected parties. In Sweden or Switzerland, special commissions, councils, or conferences were created to perpetuate consultative procedures.*

With regard to (3) evaluation, we coded:
'0': Policy effects are not comprehensively assessed, neither before nor after implementation, involving neither cost-benefit assessment nor evidence-based research.
>   Example: *The Irish public administration conducted hardly any environmental policy assessments until 2005. The rare evaluations were of limited scope and mainly driven by European Commission pressures.*

'0.5': The assessment of policy effects before or after implementation varies in scope and/or the comprehensiveness of analytical and research efforts.
>   Example: *Until Finland's EU accession, comprehensive social policy evaluations were conducted to inform policy reform, but they did not constitute a regular or formalized procedure.*

'1': Policy effects are comprehensively assessed, before and after implementation, involving cost-benefit analysis and evidence-based research.
>   Example: *In the United States, mandatory ex-ante policy evaluations are supplemented by regular ex-post impact assessments since the Governance Performance and Result Act 1993.*

instance, Belgium and Canada seem to struggle to couple policymaking and implementation in their multilayered systems of governance, other federalist countries such as Germany and Switzerland achieve rather high levels of VPI. Lastly, Figure A1 in the Supplementary material shows that in some countries and sectors, (5) values for top-down and bottom-up VPI are well aligned while in others, they develop very differently.

Alternative Explanations

Because there are factors other than VPI that might affect the burden-capacity gap, we include several control variables in our analysis. First, we control for political factors. Governments that face strong political competition have—independent of the political color of the parties in office—a particularly strong short-term incentive to meet people's demands by developing new policy targets and instruments regardless of whether they can de facto afford it given existing administrative capacities. To assess the degree of political competition, we rely on the electoral competitiveness index as provided by Kayser and Lindstädt (2015). This index estimates the "perceived loss probability" of the parties in government based on two interrelated factors: (1) whether voters will change their vote from one party to another, and (2) whether these vote shifts will ultimately make a difference for the electoral outcome, that is, the legislative seat share.

Second, to capture the effects of political institutions on the burden-capacity gap, we focus on institutional constraints and neo-corporatism. For data on institutional constraints, we rely on the indicator provided by Henisz (2002). This indicator captures not only the number of independent veto points over policy

FIGURE 4.   The Development of VPI Values across Countries, Sectors, and Time

Vertical Policy-Process Integration

Sector —— Environmental —— Social
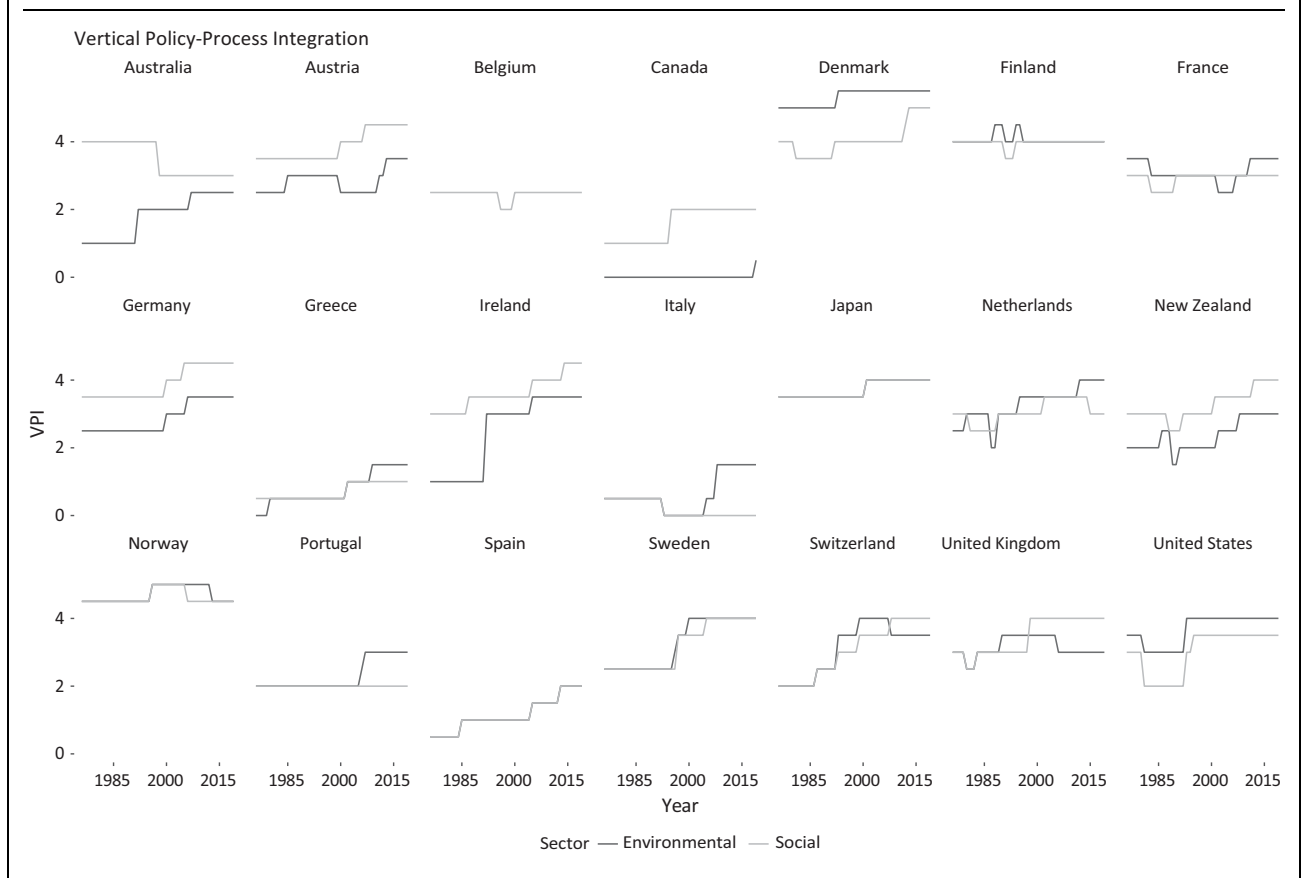
outcomes but also the distribution of preferences of the actors that inhabit them and hence whether the actors possessing veto power hold similar or divergent policy preferences. To capture the effects of corporatism, we rely on the time-variant index provided by Jahn (2016b). Jahn provides a comprehensive measurement of countries' systems of interest intermediation based on the structure, function, and scope of corporatist arrangements.

Third, we control countries' per capita GDP and the level of debt. These two variables capture the financial resources and flexibility the government possesses to either alter its policy portfolios or expand the administrative capacities available. Fourth, we check whether a country is a member of EU. EU has been a prolific producer of public policies, particularly in environmental policy. Moreover, empirical evidence suggests that member states tend to tailor supranational policies to their national context by increasing the "level of customized density" (Zhelyazkova and Thomann 2022, 439), that is, the number of policy targets and instruments in place. EU membership can hence be expected to be an additional driver of policy accumulation while not (necessarily) involving additional national administrative capacities. Lastly, countries' decisions to adopt new policies that may alter the burden-capacity gap can emerge from international policy diffusion. Here, we expect that

governments are more prone to follow one another when they are geographically close or connected via trade ties. We control these aspects by checking whether countries have a common border and by examining the share of goods being exported from one country to the other. We standardize all our continuous variables to half a standard deviation so that we can contrast their relative importance and compare continuous variables with binary ones (Gelman 2008).

EMPIRICAL FINDINGS AND DISCUSSION

In the following, we first present the results for our general model and introduce additional model specifications. Thereafter, we discuss the plausibility of our measurements and perform different robustness checks.

General Results

We explain the year-to-year changes in the ratio between the implementation burdens and the capacities by a linear model in which we control for unequal variances (heteroscedasticity, clustered errors) by country. To model time dynamics, we include an autoregressive component of order one (AR1) and control

---

BOX 1.   Model Specification

$Y_{s,c,y} \sim$	$\mathcal{N}(\mu_{s,c,y}, \sigma_{s,c,y})$	Main data component
$\mu_{s,c,y} =$	$\alpha_d + \beta_{s,v} * X_{c,t,y-1} + \theta_{s,v} * VPI_{s,c,y-3} + \rho_{s,c} * (Y_{s,c,y-1} - \mu_{s,c,y-1})$	Main linear model
$\sigma_{s,c,y} =$	$\lambda H_{c,t}$	Error component
$\alpha_d \sim$	$\mathcal{N}(0,1)$	Priors on intercepts by decade
$\beta_{s,v} \sim$	$\mathcal{MVN}(0, 2.5)$	Priors for explanatory control variables
$\theta_{s,v} \sim$	$\mathcal{MVN}(0, 2.5)$	Priors for explanatory variables on VPI
$\rho_{s,c} \sim$	$\mathcal{U}(-1, 1)$	Priors for the auto-regressive component
$\lambda_s \sim$	$\mathcal{N}(0, 10)$	Priors for heteroskedasticity controls

Where:

- s: Sector

- c: Country

- y: Year

- d: Decade

- v: Covariate

- $y_{s,c,y}$: Continuous variable with the ratio between implementation burden over implementation capacity for a specific sector (s), country (c) and year (y).

- $X_{s,c,y,v}$ Matrix with the explanatory values for each covariate (v).

- θs: Effects of covariates.

- $\rho_{s,c}$: Auto regressive component of order 1.

- σ_s: Standard deviation by sector.

- $VPI_{s,c,y,v}$ Matrix with the VPI values (two columns when two dimensions are used).

- $H_{c,y}$ Matrix with the variables controlling for heteroskedasticity (political constraints).

- α_d: Time trend. Varying intercepts by decade.

- β_s: Effects of control variables.

- θ_s: Effects of VPI on the outcome variable. Main parameters of interest.

- $\rho_{s,c}$: Auto regressive component of order 1.

---

for time-fixed effects (decade dummy). Standard errors are clustered by countries. All parameters are estimated using Bayesian inference. The full model specification can be found in Box 1. All variables are lagged by one year. For the VPI index, we use a three-year rolling average to account for the fact that political and administrative organizations and processes cannot easily be changed from one day to the next but typically take time to sediment and unfold their effects on policymaking. A three-year rolling average implies that we expect some effects from year one but that it takes up to three years until the full effect of institutional reform can be detected. We expect the strongest effect on

policymaking procedures that just started after the reforms have been completed. While policymaking procedures vary from one policy to the other and between different institutional setups (countries), we anticipate that it takes about one year for a policy to be drafted in the ministerial bureaucracy and another one to two years until the policy is discussed and ultimately passed in the parliament (Hiroi and Rennó 2018).[9]

---

[9] In section E of the Supplementary material (Figures A9, A10), we provide additional analyses with different time lags as well as with (rolling average) and without (plain lag) smoothened effects.

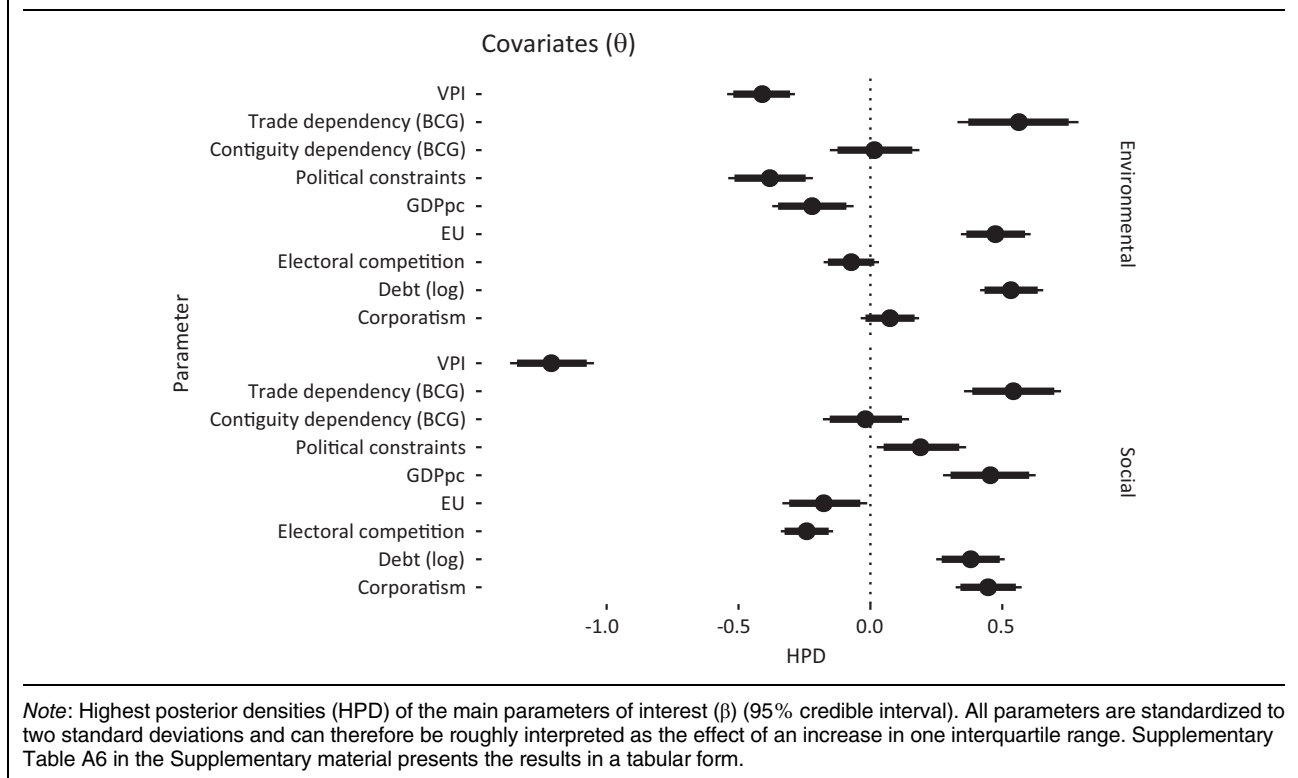**FIGURE 5.   Determinants of the Burden-Capacity Gap, Aggregated VPI (21 Countries, 1976–2018)**

Note: Highest posterior densities (HPD) of the main parameters of interest (β) (95% credible interval). All parameters are standardized to two standard deviations and can therefore be roughly interpreted as the effect of an increase in one interquartile range. Supplementary Table A6 in the Supplementary material presents the results in a tabular form.

FIGURE 6.   Magnitude of the VPI Effect: Expected Change in the Burden-Capacity Gap

Note: The gray areas define 95% credible intervals. Descriptive statistics are shown in Supplementary Table A5, with full model specifications in Supplementary Table A6.
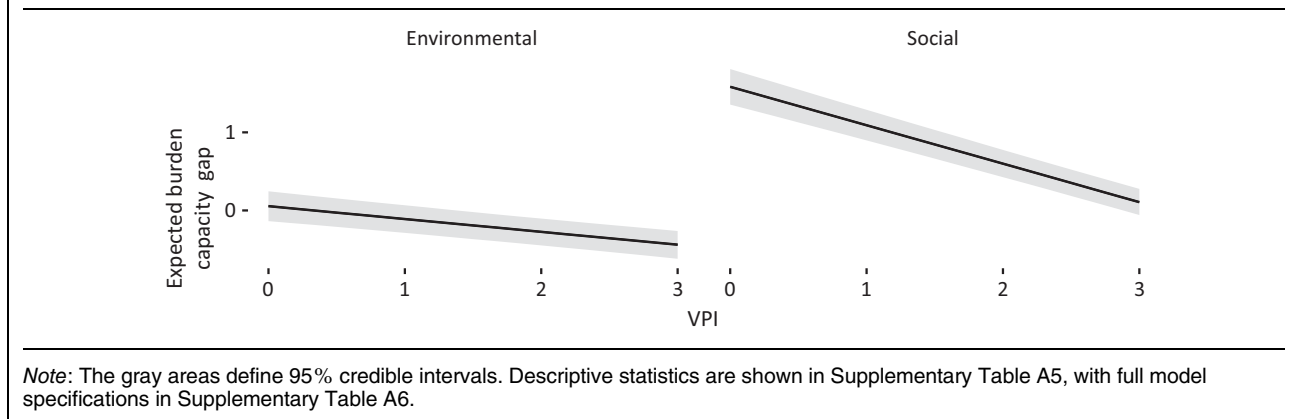

Figure 5 presents our key results. They can also be looked up in section D of the Supplementary material, together with additional material on our model testing. In total, the factors considered account for 59% of the variance of the dependent variables in environmental and 56% in social policy. The empirical analysis reveals that higher levels of VPI lead to a stronger match between policy accumulation and the administrative capacities available in both areas.

In Figure A17 of the Supplementary material, we provide an additional analysis, separately assessing the impact of VPI on the numerator and denominator. When examining the unique impact of VPI on each

"side" of the gap, our findings reveal that higher levels of VPI consistently correlate with fewer policies, yet greater capacities in social policy. However, in the context of environmental policy, VPI principally results in an increased allocation of resources for implementers, rather than a reduction in the number of policies.

This outcome appears logical given that environmental policy is frequently depicted as a highly internationalized policy field, particularly in areas such as air and water quality, which were the focus of our study (Knill, Debus, and Heichel 2010). Despite the perception of reduced leeway for governments to produce fewer policies within this field, our research indicates that

enhanced VPI still results in a smaller gap size. This suggests two primary points: first, the burden-capacity gap is a distinct empirical phenomenon that is not merely about having more or fewer policies and it represents a different consideration altogether, and second, it needs the VPI to comprehensively understand this phenomenon.

But how strong is the VPI effect exactly? Figure 6 provides a detailed assessment of the magnitude of the VPI's effects when all other variables are kept at their mean values (see Supplementary Table A5). In environmental policy, moving from the minimum to the maximum VPI value implies a decrease in the burden-capacity gap from 0.10 to –0.40. This sounds rather small but equals a change of almost half a standard deviation. In the area of social policy, the effect is even more pronounced; here, the gap size decreases from an expected 1.6 to 0.10 when the VPI goes from the minimum to the maximum value. This equals a reduction of almost one and a half standard deviations.

In contrast to the degree of VPI, the effects of most other variables vary across the policy sectors—not only in terms of their strength but also in terms of direction. These differing effects, however, are quite plausible when considering the peculiarities of the respective policy areas. First, it is for social policy only that we observe the effects of electoral competition on the burden-capacity gap. Here, higher competition contributes to a lower gap. As voters are directly affected by social policies and are hence highly aware of these measures' effectiveness, merely demonstrating responsiveness might not be sufficient for vote-seeking politicians. They need to ensure that these policies actually work (for comparable findings, see Dasgupta and Kapur 2020).

Second, political constraints seem to restrict excessive growth in environmental policy but have the exact opposite effect on social policy. An explanation could be that the redistributive character of social policies requires governments to engage in side payments and package deals when consensus requirements are particularly high (Scharpf 1988). This argument also explains that corporatism is associated with a higher burden-capacity gap in social policy. Close cooperation between government and societal associations seems to facilitate the adoption of policy packages at the expense of third parties not participating in these negotiations (here the implementation level).

Third, higher levels of debt lead to a greater mismatch between policy accumulation and available administrative capacities in both policy areas. Essentially, this implies that governments facing budgetary pressure do not tend to produce less policies but in fact, save on administrative capacities. This finding resonates with the argument made above that policymakers face perverse incentives to overproduce policies even when knowing that effective implementation cannot be guaranteed. For the level of economic prosperity, we find the exact opposite relationship.

Fourth, turning to international factors, EU membership tends to increase the burden-capacity gap in environmental policy but to have an inverse effect on social policy. This is highly plausible, as the EU possesses far-reaching policymaking competencies in environmental matters. By contrast, member states have transferred considerably less social policy competences to the EU[10]. In addition, diffusion effects emerging from higher trade dependence also come with higher burden-capacity gaps in both sectors. Close trade connections seem to stimulate the adoption of additional policies. The most likely reasons for this connection are the harmonization of domestic regulatory standards in the case of environmental policy (Vogel 1995) and governments' attempts to compensate globalization losers for the risks associated with increased exposure to international trade in the case of social policy (Walter 2010). It seems that this trade-induced policy growth is not compensated by corresponding expansions of administrative capacities.

Up until this point, our analysis has primarily focused on examining how the formation of VPI has influenced the burden-capacity gap. To explore what happens when we disaggregate the concept into its constituent dimensions, Figure 7 includes the bottom-up and top-down VPI as separate variables. The findings show that both dimensions exhibit a negative and significant influence on the burden-capacity gap. These findings empirically support our logic of concept formation, specifically the notion of substitutability.

## Plausibility and Robustness Checks

There are naturally several aspects of our conceptual and methodological approach that require further discussion. First, our analysis relies on macro-level concepts to size the burden-capacity gap. The accuracy of the empirical findings thus strongly depends on whether our aggregate estimates are able to quantitatively capture real-world differences. To check this aspect, we provide a "proof of concept" in the Supplementary material (section F, Figure A18). We demonstrate that the policy effectiveness of (environmental) policies generally decreases with higher gap sizes and that new measures become ineffective at a certain tipping point.

Moreover, to cross-validate and check the plausibility of our central measures, we conducted 36 semistructured in-depth interviews with public employees responsible for implementing social and environmental policies in Denmark and Italy.[11] In line with Lieberman's (2005) suggestion of a "nested analysis," we chose Italy and Denmark based on our previous large-N study. These two countries present the most pronounced differences regarding the burden-capacity

---

[10] In the Supplementary material (E.8), we included the Regional Authority Index (Hooghe et al. 2016) as an additional control accounting for multilayered policymaking systems. Our results remain unchanged.

[11] Human subjects research in this article was reviewed and approved by the Ethics Commission of the Social Science Faculty of LMU Munich (Certificate No. GZ 18-01). Section G of the Supplementary material describes the sampling process, the interview strategy, and the detailed results of the coding procedure.

FIGURE 7. Determinants of the Burden-Capacity Gap, Disaggregated VPI (21 Countries, 1976–2018)
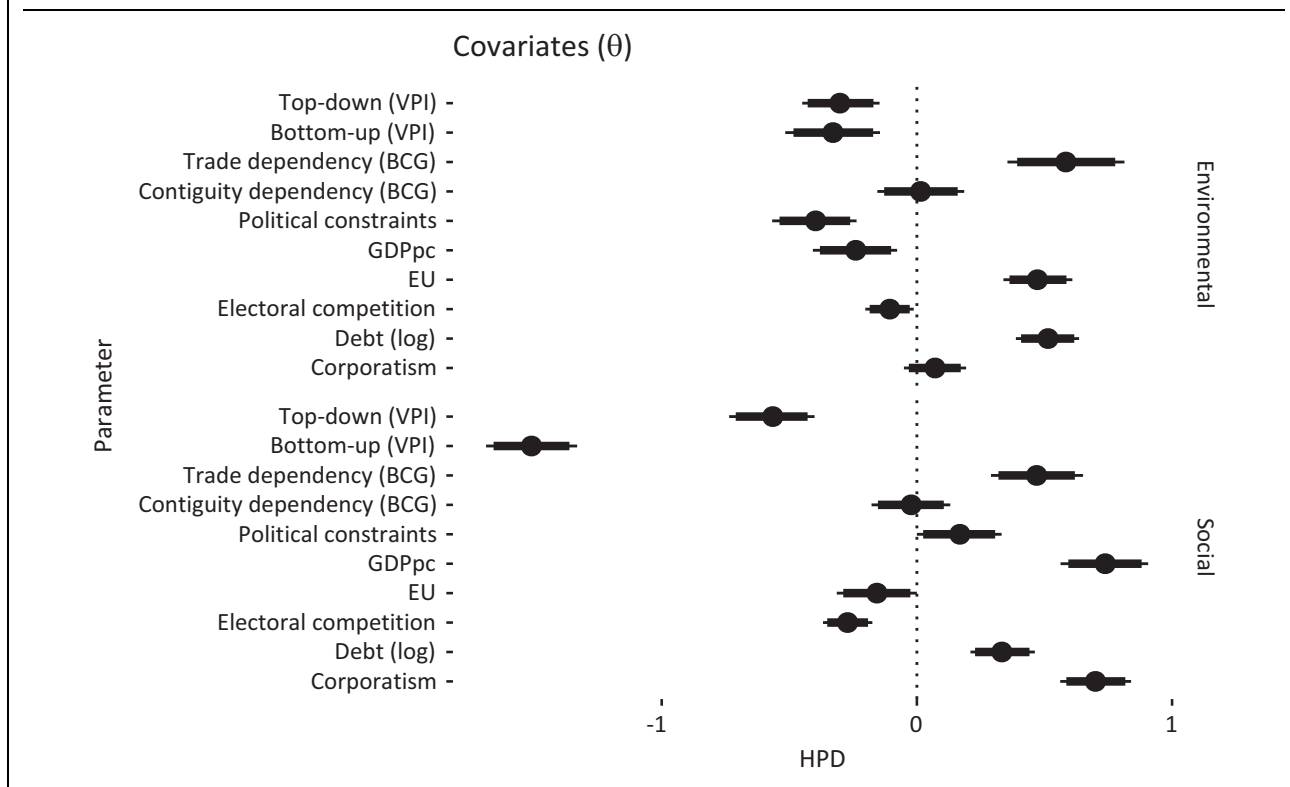
Note: HPD of the main parameters of interest (β) (95% credible interval). All parameters are standardized to two standard deviations and, therefore, can be roughly interpreted as the effect of an increase in one interquartile range. Supplementary Table A7 presents the results in a tabular form.

gap (see Figure 3). While the gap is particularly large in the Italian case in both sectors, Denmark effectively succeeds in keeping a due balance between implementation burdens and the administrative capacities available. If our quantitative measurements are correct, these marked differences should also be reflected in the work experiences of public administrators.

The qualitative insights from the interviews strongly confirm the impression gained from our aggregate measures. In Denmark, the policy implementers reported that the provided resources strongly match the capacities necessary for adequate implementation. They indicate that they "don't see a lot of performance difficulties,"[12] "have enough scope for pursuing [their] tasks," and "are not asked anything that is impossible for [them] to deliver." In Italy, the picture is exactly the opposite. In both environmental and social policy, there are numerous accounts of "overburdened" and "struggling" implementing authorities that find themselves "in the eye of a storm" as "work (…) rains down on [them] from above." Implementers state that "[resource] constraints affect (…) the impact of policies

and (…) the achievement of objectives" and that, in practice, there is a "big disconnect between the political and the technical [implementation] side."

A second important assumption in our quantitative analysis is that all policies imply more or less the same implementation burden for the administration. In reality, however, it might be the case that policies strongly differ in what they require from the administration. While it is difficult to give varying weights to different policies from the outset, we can expect that over time, the administration might benefit from learning effects, that is, the acquisition of knowledge about the functioning of policies and the required implementation processes. In other words, administrators should find it easier to implement the types of policies they already know (Fernández-i-Marín et al. 2023b).

To account for such learning effects, we model different ideal-type learning curves. More precisely, we discount increases in the policy portfolios based on how extensively the government has previously used a given instrument type. As illustrated in Figure 8, the "no-learning" scenario represents our standard approach where all additional target-instrument combinations result in one additional "unit" of implementation burden (1; 1; 1; 1). In the "continuous learning" scenario, only the first target addressed by a given instrument is fully counted and gets the value of 1. The

---

[12] The direct quotes in this section are statements that are representative of the views of several interviewees. See Table A27 in Section G of the Supplementary material for further details and quotes.

FIGURE 8.   Learning Effects on the Relationship between Policy Growth and Implementation Burdens
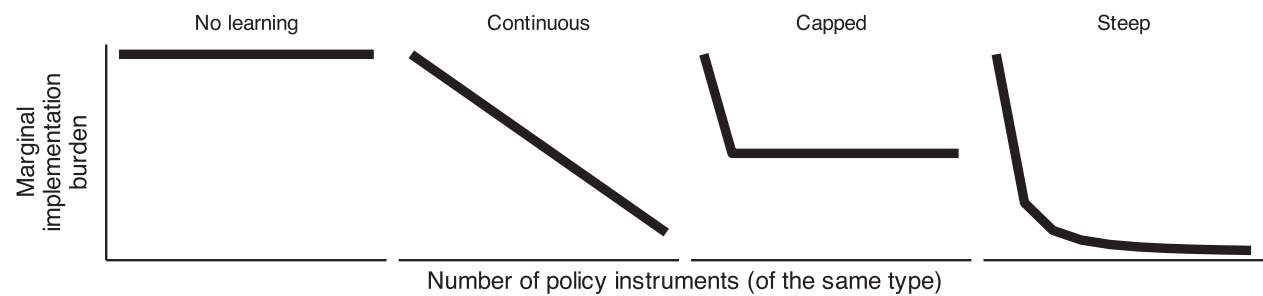
Note: Different curves represent the marginal implementation burdens caused by additional policies under the assumption of different learning curves of implementing authorities.

FIGURE 9.   Determinants of the Burden-Capacity Gap, Aggregated VPI (21 Countries, 1976–2018) with Learning Effects
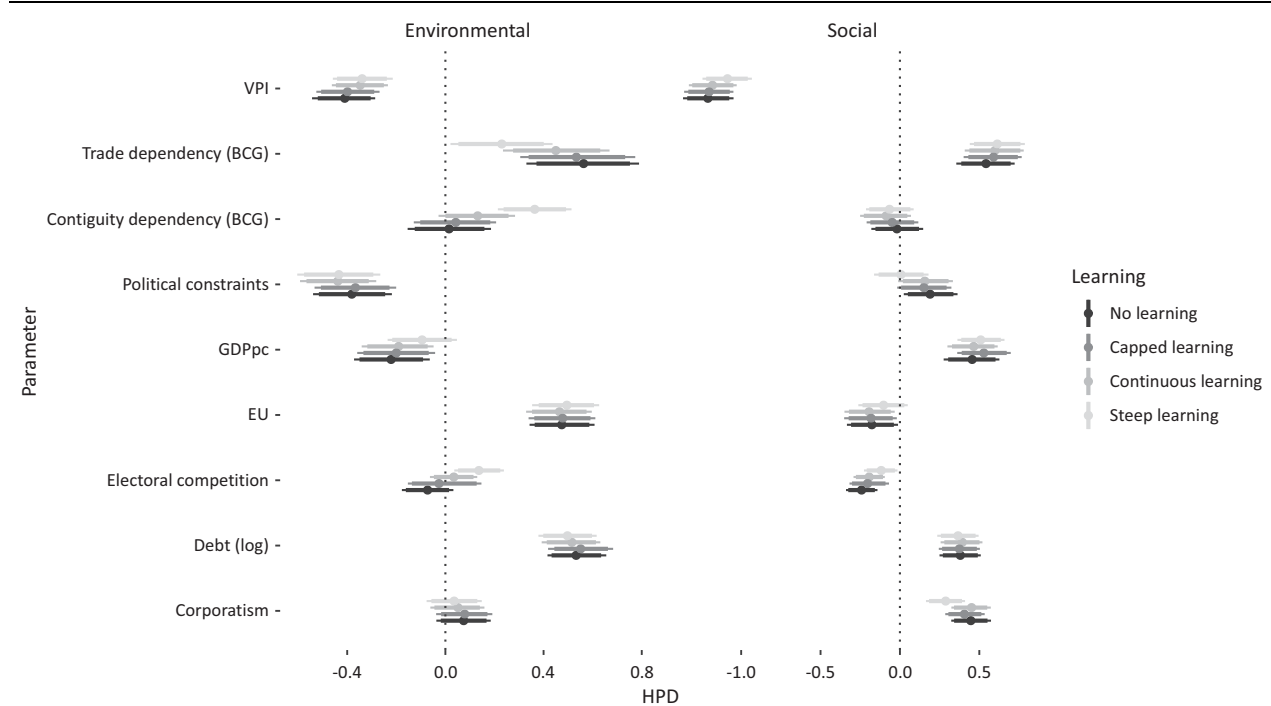
Note: HPD of the main parameters of interest (β) (95% credible interval). All parameters are standardized to two standard deviations. Supplementary Tables A6 (no learning) and A8 to A10 (learning) present the results in a tabular form.

remaining policy targets addressed by the *same* instrument type get shares of the original implementation burden load $(1; 1/2; 1/4; 1/8$, etc.). The "capped learning" scenario, by contrast, implies a one-time difference in the marginal implementation burden. After this initial decrease, the additional implementation burden induced by portfolio expansions stays constant $(1; 1/2; 1/2; 1/2$, etc.). Lastly, "steep learning" describes a scenario in which the marginal implementation burdens diminishes rapidly with each additional policy $(1; 1/4; 1/9; 1/16$, etc.).

In Figure 9, we present our results when we replicate our previous analysis (see again Figure 4) while assuming different learning scenarios that mediate the impact of policy accumulation on implementation burdens.[13] While the impact of the VPI (very)

---

[13] In section E.6 of the Supplementary material, we replicate this analysis, putting less weight on instruments that fall on the same policy target (learning via targets). Here, the underlying logic is that administrations will find it easier to implement policies once they

slightly decreases in size, our central findings remain unchanged.

Lastly, a challenge for our analysis might be that there is a negative correlation between VPI and the burden-capacity gap as both sides of the equation should benefit from higher levels of *general* state capacity. From this perspective, our analysis might suffer from an omitted variable bias. To address this concern, section E.7 of the Supplementary material (Figure A15) includes general state capacity as provided by Hanson and Sigman (2021) as an additional control. The analysis reveals that VPI has its own distinct effect on the burden-capacity gap, and this persist even when controlling for general state capacity as a potential confounder.

CONCLUSION

We started this paper with the observation that democratic governments tend to produce more and more policies, thus putting an ever-greater burden on the implementation level. Political responsiveness to societal demands comes with a continuous growth of policies. At the same time, policy growth leads to an accumulation of burdens for the bodies in charge of implementing these policies. If administrative capacities do not keep pace with ever-growing implementation burdens, this will undermine policy effectiveness, hence the long-term legitimacy of modern democracies.

In our analysis, we offer a novel approach to systematically measure and explain the development of the burden-capacity gap across countries and sectors and over time. We showed that national and sectoral variations of this gap are strongly affected by the extent to which processes of policy formulation and implementation are vertically integrated. Higher levels of VPI are strongly associated with a smaller burden-capacity gap. The existence of institutional arrangements that link the bureaucracies in charge of policymaking and those involved in policy implementation thus constitutes a crucial feature of bureaucratic quality that has so far been fairly neglected in academic debates. In fact, VPI is the only factor considered in our analysis (aside from a country's economic wealth) that consistently and effectively affects the growth of the gap between implementation burdens and administrative capacities.

While VPI has a very strong and pronounced effect in this regard, a more fine-graded assessment of different VPI dimensions reveals that the effects of bottom-up integration tend to be even more pronounced than the effects of top-down integration. From a practitioner's perspective, this is a promising finding. While the (re-)allocation of costs and competencies typically involves a strong political conflict and is thus difficult to pursue, the inclusion of policy implementers'

---

have gained more knowledge about the policy targets (as opposed to the functioning of the same policy instruments). Again, our central results are not altered.

experience in policy formulation can be achieved relatively easily, through the creation and reform of both formal and informal institutions. Governments thus have potent tools at their disposal that if applied could safeguard democracies from overloading themselves with ineffective policies.

A limitation of our research is that in measuring policy growth, we focused on the introduction of new policy targets and instruments. Yet, the implementation burden might also increase in response to changes in the calibration of *existing* policies (derogations, exemptions, special clauses, etc.) which might affect the complexity, thus the time and effort that must be invested in applying a given policy provision. These increases in implementation burdens that emerge beyond changes in the mere number of policies constitute an important area for future research.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0003055423001090.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/ZNRTYA.

ACKNOWLEDGMENTS

FUNDING STATEMENT

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

ETHICAL STANDARDS

The authors declare the human subjects research in this article was reviewed and approved by the Ethics Commission of the Social Science Faculty of LMU Munich and certificate numbers are provided in the text (see footnote 11). The authors affirm that this article adheres to the APSA's Principles and Guidance on Human Subject Research.

REFERENCES

Adam, Christian, Steffen Hurka, Christoph Knill, and Yves Steinebach. 2019. *Policy Accumulation and the Democratic Responsiveness Trap*. Cambridge: Cambridge University Press.

Andersen, David D. E. 2018. "Does Meritocracy Lead to Bureaucratic Quality? Revisiting the Experience of Prussia and Imperial and Weimar Germany." *Social Science History* 42 (2): 245–68.

Ansell, Chris, and Alison Gash. 2008. "Collaborative Governance in Theory and Practice." *Journal of Public Administration Research and Theory* 18 (4): 543–71.

Bäck, Hanna, and Axel Hadenius. 2008. "Democracy and State Capacity: Exploring a J-Shaped Relationship." *Governance* 21 (1): 1–24.

Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. 2020. "The Lay of the Land: Information Capacity and the Modern State." *Comparative Political Studies* 53 (2): 175–213.

Cohen, Nissim. 2021. *Policy Entrepreneurship at the Street Level: Understanding the Effect of the Individual*. Cambridge: Cambridge University Press.

Cohen, Nissim, and Neomi Frisch-Aviram. 2021. "Street-Level Bureaucrats and Policy Entrepreneurship: When Implementers Challenge Policy Design." *Public Administration* 99 (3): 427–38.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, *et al.* 2019. V-Dem [Country-Year/Country-Date] Dataset v9. Varieties of Democracy (V-Dem) Project. doi:10.23696/vdemcy19.

Dahlberg, Stefan, and Sören Holmberg. 2014. "Democracy and Bureaucracy: How Their Quality Matters for Popular Satisfaction." *West European Politics* 37 (3): 515–37.

Dahlström, Carl, Victor Lapuente, and Jan Teorell. 2012. "The Merit of Meritocratization: Politics, Bureaucracy, and the Institutional Deterrents of Corruption." *Political Research Quarterly* 65 (3): 656–68.

Dasgupta, Aditya, and Devesh Kapur. 2020. "The Political Economy of Bureaucratic Overload: Evidence from Rural Development Officials in India." *American Political Science Review* 114 (4): 1316–34.

Fernández-i-Marín, Xavier, Markus Hinterleitner, Christoph Knill, and Yves Steinebach. 2023a. "Bureaucratic Overburdening in Advanced Democracies." *Public Administration Review*. https://doi.org/10.1111/puar.13725.

Fernández-i-Marín, Xavier, Markus Hinterleitner, Christoph Knill, and Yves Steinebach. 2023b. "Policy Growth, Implementation Capacities, and the Effect on Policy Performance." *Governance*. https://doi.org/10.1111/gove.12816.

Fernández-i-Marín, Xavier, Christoph Knill, Christina Steinbacher, and Yves Steinebach. 2023. "Replication Data for: Bureaucratic Quality and the Gap between Implementation Burden and Administrative Capacities." Harvard Dataverse. Dataset. https://doi.org/10.7910/DVN/ZNRTYA.

Fukuyama, Francis. 2013. "What is Governance?" *Governance* 26 (3): 347–68.

Gelman, Andrew. 2008. "Scaling Regression Inputs by Dividing by Two Standard Deviations." *Statistics in Medicine* 27 (15): 2865–73.

Gemeindetag Baden-Württemberg. 2022. "Belastungsgrenze überschritten. Es darf kein „Weiter so" geben."

Gerring, John. 2008. "Case Selection for Case-Study Analysis: Qualitative and Quantitative Techniques." In *The Oxford Handbook of Political Methodology*, eds. Janet M. Box-Steffensmeier, Henry E. Brady and David Collier, 645–84. Oxford: Oxford University Press.

Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.

Gratton, Gabriele, Luigi Guiso, Claudio Michelacci, and Massimo Morelli. 2021. "From Weber to Kafka: Political Instability and the Overproduction of Laws." *American Economic Review* 111 (9): 2964–3003.

Hanson, Jonathan K., and Rachel Sigman. 2021. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." *The Journal of Politics* 83 (4): 1495–510.

Healy, Andrew, and Neil Malhotra. 2013. "Retrospective Voting Reconsidered." *Annual Review of Political Science* 16: 285–306.

Henisz, Witold J. 2002. "The Institutional Environment for Infrastructure Investment." *Industrial and Corporate Change* 11 (2): 355–89.

Herd, Pamela, and Donald P. Moynihan. 2018. *Administrative Burden – Policymaking by Other Means*. New York: Russell Sage Foundation.

Hill, Michael, and Peter Hupe. 2009. *Implementing Public Policy: An Introduction to the Study of Operational Governance*. London: SAGE.

Hinterleitner, Markus. 2020. *Policy Controversies and Political Blame Games*. New York: Cambridge University Press.

Hiroi, Taeko, and Lucio Rennó. 2018. "Disentangling Legislative Duration in Coalitional Presidential Systems." *The Journal of Legislative Studies* 24 (4): 475–98.

Hooghe, Liesbet, and Gary Marks. 2003. "Unraveling the Central State, but How? Types of Multi-level Governance." *American Political Science Review* 97 (2): 233–43.

Hooghe, Liesbet, Gary Marks, and Arjan H. Schakel. 2008. "Operationalizing Regional Authority: A Coding Scheme for 42 Countries, 1950–2006." *Regional & Federal Studies* 18 (2-3): 123–42.

Hooghe, Liesbet, Gary Marks, Arjan H. Schakel, Sandra Chapman Osterkatz, Sara Niedzwiecki, and Sarah Shair-Rosenfield. 2016. *Measuring Regional Authority: A Postfunctionalist Theory of Governance*. Oxford: Oxford University Press.

Howlett, Michael, and M. Ramesh. 2016. "Achilles' Heels of Governance: Critical Capacity Deficits and Their Role in Governance Failures." *Regulation & Governance* 10 (4): 301–13.

Hurka, Steffen, Maximilian Haag, and Constantin Kaplaner. 2022. "Policy Complexity in the European Union, 1993-Today: Introducing the EUPLEX Dataset." *Journal of European Public Policy* 29 (9): 1512–27.

Jahn, Detlef. 2016a. *The Politics of Environmental Performance*. Cambridge: Cambridge University Press.

Jahn, Detlef. 2016b. "Changing of the Guard: Trends in Corporatist Arrangements in 42 Highly Industrialized Societies from 1960 to 2010." *Socio-Economic Review* 14 (1): 47–71.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2011. "The Worldwide Governance Indicators: Methodology and Analytical Issues." *Hague Journal of the Rule of Law* 3 (2): 220–46.

Kayser, Mark A., and René Lindstädt. 2015. "A Cross-National Measure of Electoral Competitiveness." *Political Analysis* 23 (2): 242–53.

Klein, Ezra. 2022. "Opinion: What America Needs Is a Liberalism That Builds." *New York Times*, May 29. https://www.nytimes.com/2022/05/29/opinion/biden-liberalism-infrastructure-building.html.

Knill, Christoph, Marc Debus, and Stephan Heichel. 2010. "Do Parties Matter in Internationalised Policy Areas? The Impact of Political Parties on Environmental Policy Outputs in 18 OECD Countries, 1970–2000." *European Journal of Political Research* 49 (3): 301–36.

Knill, Christoph, Christina Steinbacher, and Yves Steinebach. 2021. "Sustaining Statehood: A Comparative Analysis of Vertical Policy-Process Integration in Denmark and Italy." *Public Administration* 99 (4): 758–74.

Knill, Christoph, Yves Steinebach, and Dionys Zink. 2023. "How Policy Growth Affects Policy Implementation: Bureaucratic Overload and Policy Triage." *Journal of European Public Policy*. https://doi.org/10.1080/13501763.2022.2158208.

Lieberman, Evan S. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99 (3): 435–52.

Limberg, Julian, Yves Steinebach, Louisa Bayerlein, and Christoph Knill. 2021. "The More the Better? Rule Growth and Policy Impact from a Macro Perspective." *European Journal of Political Research* 60 (2): 438–54.

Linder, Stephen H., and B. Guy Peters. 1984. "From Social Theory to Policy Design." *Journal of Public Policy* 4 (3): 237–59.

Lodge, Martin, and Kai Wegrich. 2014. *The Problem-Solving Capacity of the Modern State: Governance Challenges and Administrative Capacities*. Oxford: Oxford University Press.

Moffitt, Susan L., Cadence Willse, Kelly B. Smith, and David K. Cohen. 2021. "Centralization and Subnational Capacity: The Struggle to Make Federalism Work Equitably in Public Education." *Perspectives on Politics*. https://doi.org/10.1017/S1537592721002012.

Nicholson-Crotty, Jill, and Susan M. Miller. 2012. "Bureaucratic Effectiveness and Influence in the Legislature." *Journal of Public Administration Research and Theory* 22 (2): 347–71.

OECD. 2020. "Social Expenditure Database (SOCX)." OECD Statistics. https://www.oecd.org/social/expenditure.htm.

Peters, B. Guy, and Jon Pierre. 2004. "Multi-Level Governance and Democracy: A Faustian Bargain?" In *Multi-Level Governance*, eds. Ian Bache and Matthew Flinders, 75–90. Oxford: Oxford University Press.

Pressman, Jeffrey L., and Aaron Wildavsky. 1973. *Implementation: How Great Expectations in Washington are Dashed in Oakland; Or, Why It's Amazing that Federal Programs Work at All*. Berkeley: University of California Press.

Rauch, James E., and Peter B. Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries." *Journal of Public Economics* 75 (1): 49–71.

Reuning, Kevin, Michael R. Kenwick, and Christopher J. Fariss. 2019. "Exploring the Dynamics of Latent Variable Models." *Political Analysis* 27 (4): 503–17.

Sager, Fritz, and Anat Gofen. 2022. "The Polity of Implementation: Organizational and Institutional Arrangements in Policy Implementation." *Governance* 35 (2): 347–64.

Scharpf, Fritz W. 1988. "The Joint-Decision Trap: Lessons from German Federalism and European Integration." *Public Administration* 66 (3): 239–78.

Schneider, Anne L., and Helen M. Ingram. 1997. *Policy Design for Democracy*. Lawrence: University Press of Kansas.

Stockmann, Reinhard, Wolfgang Meyer, and Lena Taube. 2020. *The Institutionalization of Evaluation in Europe*. Cham, Switzerland: Palgrave MacMillan.

Vogel, David. 1995. *Trading Up: Consumer and Environmental Regulation in a Global Economy*. Cambridge, MA: Harvard University Press.

Walter, Stefanie. 2010. "Globalization and the Welfare State: Testing the Microfoundations of the Compensation Hypothesis." *International Studies Quarterly* 54 (2): 403–26.

Winter, Søren C. 2012. "Implementation Perspectives: Status and Reconsideration." In *Handbook of Public Administration*, eds. B. Guy Peters and Jon Pierre, 265–78. London: SAGE.

World Bank. 2023. "World Development Indicators." https://databank.worldbank.org/source/world-development-indicators.

Zhelyazkova, Asya, and Eva Thomann. 2022. "'I Did it My Way': Customisation and Practical Compliance with EU Policies." *Journal of European Public Policy* 29 (3): 427–47.