

RESEARCH WORKSHOP ON "POLICY FAILURES"  
NATIONAL UNIVERSITY OF SINGAPORE  
February, 20-21 2014

# Defining Recipes for Policy Failures

---

## A Realist, Configurational Approach

Alessia Damonte

Dept. of Social and Political Sciences  
Università degli Studi di Milano  
[alessia.damonte@unimi.it](mailto:alessia.damonte@unimi.it)

-- first draft --

**Abstract.** The article reasons that the tenets around which the analyses of policy failures often revolve may lead to flawed causal conclusions. Building on data about policy tools for green growth, it hence demonstrates how Qualitative Comparative Analysis can overcome these shortcomings if applied to a Realist understanding of poor policy results – thus contributing to the debate on models and strategies for explaining policy failures.

## Introduction

As many concepts in social sciences, that of 'policy failure' is only apparently straightforward: when boiled down to a precise analytical model, much of its boundaries and contents depend on the literature one refers to. Despite differences, however, definitions seem to agree on three intertwined tenets: that failure is a linear yet multidimensional opposite of success; that its understanding mainly rests on remote antecedents; and that explanations are necessarily overdetermined (Howlett 2012; McConnell 2010, 2011; Bovens, 't Hart and Peters 1998, 2001; Gray 1998, 1996; Bovens and 't Hart 1996).

These three tenets make perfect sense within an interpretivist strategy: they equip researchers for hunting crucial sequences with an encompassing concept of failure and a preference for case studies. Yet, such equipment proves unsuitable for explanatory research: studies committed to these tenets could neither demonstrate crucial sequences to be actual causes of failure nor that success would have followed had these sequences been avoided.

To the extent that a clear ascription of causality is essential to blame avoidance and lesson learning, an analytical shift is then required that makes unambiguous explanation possible. This article aims to contribute to failure studies by indicating a reasoned ontological and methodological direction for it.

After examining the limits of the three tenets (§1) the article thus considers the gains of adopting a 'Realist' approach to policy failings (§2), hence of applying the Qualitative Comparative Analysis to explanatory models (§3). The issue of green growth in the EU15 is then presented as a case for the joint exercise of the analytical approach and of the method (§4). Conclusions (§5) will discuss results.

### **1. Three tenets in policy failure studies and reasons to shift away when looking for causality**

In reviewing the failure literature, Howlett remarks how a «common way to treat policy failure [...] is as the reverse of policy success in the sense that whatever does not succeed is a failure» (2012:542). At the same time, this literature also acknowledges that policies are complex, so that they can crash along many dimensions. In the effort of providing more accurate accounts of such multidimensionality, definitions have thus increasingly widened over time.

While Dunleavy (1995:52) considered «disasters» as «significant and substantially costly» decision or non-decisions, Gray (1996:74–83) confined them to

those policies alone which had failed against «nearly every possible» criteria – so implying that failures occur whenever some relevant standard is not met. In addressing the issue from the side of policy success, Moore (1995:71) specified these relevant criteria in (a) legitimation, (b) capacity to deliver the intended goods to the right beneficiaries, (c) satisfaction of main clients and stakeholders; the lack of even one property was deemed enough for a policy to fail (McConnell 2011). The relevance of social discontent was especially emphasized by Bovens and 't Hart, who focused on «fiasco» as «a negative event perceived by a socially and politically significant group of people in the community to be at least partially caused by avoidable and blameworthy failures of public policymakers» (1996:15). Fiascoes are hence different from failures in that the former impute responsibilities independently from evaluators' findings, and are politicized. Later, Bovens, 't Hart and Peters singled «program failure», which «occurs when a policy decision, plan or strategy that has been implemented fails to have the desired impact», out of «political failure», indicating a lack of «the support and momentum necessary for its (long-term) survival as a prioritized area of government activity» (1998:200). A fourfold typology was then derived (2001) that McConnell (2010b) later expanded by adding a third analytical layer. After him, policies can be classified as failures or successes in each dimension according to special 'indicators': whether the expected goods are delivered to the intended recipients (program); how legitimate the decision-making is deemed and whether it consolidates coalitions (process); how the policy strengthens the decision-makers' reputation and prospects (political) (*ibid.*:46). Howlett (2012) further specified the properties of a failure in its extent, duration, visibility, avoidability, intentionality, intensity; and supplemented the framework with deductive prescriptions.

The search for higher definitional precision has thus stretched the concept to cover the many failings a policy can undergo, and constructed types by combining kinds of failures. Such extension is justified from the epistemic perspective of case-oriented analyses, where expectations and prescriptions can follow if the nature of a phenomenon is understood exactly within an «explanatory typology» (Elman 2005). However, from the viewpoint of causal analysis, such extension comes at some price – heuristic, logical, normative and ontological.

Heuristically, the classification implies the tautological identification of the *explanandum* and of the *explanans* as far as it is used not to develop testable expectations of the types' behavior, but to construct the object of analysis while accounting for it. Logically, as full success depends on the joint non-occurrence of the many possible failings, the framework states the practical impossibility of policy success,

and legitimizes a focus on negative cases only that can bias conclusions (Most and Starr 2003). Normatively, in establishing low consensus as a kind of failure, the framework assumes as possible and desirable a world where policies can please the whole of the community interests – somehow at odds with the basic democratic requisite of interest pluralism; moreover, in setting the survival of a program as an indicator of success, it takes a questionable stance in favor of the stability of policy regimes over time. Finally, the framework treats the dimensions as ontologically equal criteria, so dismissing the causal relationship between them as either irrelevant or given – again a presumption that may rather deserve an empirical assessment.

The ontology implied by the extended concept of failure is usually completed by another assumption: that government is the chief cause, hence what the research shall focus on for prescriptions.

Such assumption may result from a deliberate shortcut through an intricate empirical representation: Bovens and 't Hart's analysis highlighted how disputable the identification is of crucial actors and key sequences, hence the ascription of responsibilities (1996:60); later, they explicitly decided to narrow the list of candidates for blame down to governments alone – as, despite «... not having control over many of the factors that may have produced the problems», «governments bear the ultimate burden for success and, most of all, failure» (Bovens, 't Hart and Peters 2001:12; Howlett 2012). Also, when policy failure is directly framed from the perspective of the rulers' interest and strategy, a government-centered causation may turn out to be implied by the definition itself: for instance in McConnell (2010b), where the success of a program «is a more-or-less interrupted achievement of what government set out to achieve» (*ibid.*:67); a successful process «basically 'gets the job done' for government» (*ibid.*:64); and political success overlaps «the classic 'vote-winning' or reputation-building policy» (*ibid.*:73). Whatever the reason, such analytical framework embodies the time-honored belief that politics determines policy – and, given the delegation chain, as any performance cannot but follow the properties of the top, the top is the ultimate antecedent of any critical sequence – and the only relevant to the analysis.

Such bias may have paradoxical consequences on the usability of research findings. Even discounting the problems of ambiguity that structurally affects governmental policy decisions (Matland 1995) and of the many 'democratic deficits' that blur the chain of delegation (Fung 2006), to some scholars governments are unescapably doomed to fail by design: either because they concentrate too much power in their departments, so that any normative preference can be implemented regardless

of how reasonable and justifiable it actually is (Moran 2001); or because they are too weak, and hence some minorities are «granted concessions ... in the design of public bureaucracies whose policy mandates they oppose» which «allow them to weigh these organizations down with wholly unsuitable structures that promote ineffectiveness and failure» (Moe 1995:126-7). Under these conditions, to avoid government failure would simply become impossible, and research useless.

Moreover, a focus on antecedents implies but does not prove that special sequences would necessarily follow certain governmental properties; also, it may lead to neglect some «proximate» factors despite their recognized relevance as «efficient» causes – i.e., the capture of policies by special internal and interests (Niskanen 1971; Olson 1982; Carlsson 2000); the «'war of attrition' amongst interest groups [...] hoping that another group will bear the costs» (Lundahl and Wyzan 2005:16); and misfitting policy instruments (Lundahl and Wyzan 2005; Vining & Weimar 1990; Wolf 1987; King 1975).

A focus on government as major cause of policy failures may then result into omissive explanatory models, and useless findings.

Besides, that of causality does not seem a conventional concern in failure literature. Many relevant studies imply a commitment to an interpretative epistemology that pairs with a conscious dismissal of causality as a research goal in favor of 'how questions' alone. Remarkably explicit were Bovens and 't Hart, who wrote of their work: «it is not a systematic empirical study [...] designed to uncover perhaps not timeless proverbs but contingent generalizations about how and why things go wrong in public policymaking. What this study seeks to do is to shed light on the ways in which people and societies come to terms with highly negative events, and to what extent and how they associate this with the actions and policies of their governments» (1996:18).

Such dismissal of standard explanatory strategies was justified by the recognition that distinct decisive interests evaluate the very same result quite differently – which downplays the relevance of 'the men with the method' and of their evaluation. Furthermore, the resulting multiplicity of legitimate opinions has been placed within a dynamic portrait of the policy process; possible causation so overlaps meaningful sequences in a muddle of relentless and sometimes whimsy accommodations of means to multiple changing policy ends. The analytical exercises used to make sense out of such 'complicatedness' – description, interpretation, classification – rest on the researchers' informed appreciation of within-case special 'co-variations' of occurrences. Yet, researchers' appreciation heavily depends on how evidence is sifted by the special

theory in use, whose explanatory power is rather assumed than verified. As almost any theory can find some supporting evidence from any case, these accounts often result into syncretic compositions of theories and the conclusion that «one fundamental feature of disasters is precisely the existence of multiple causes: an 'overdetermination', in which any number of potential causes could singly – or in combination – lie behind débacle» (Gray 1998:7).

This uncertain causality however depends less on the variety of available theories than on a preference for in-depth research strategies – which failure analyses also maintain when they explicitly aim to compare and explain. Bovens, 't Hart and Peters (2001) did provide 24 detailed case studies, from 6 countries selected because of their combination of institutional and political characteristics, in order to verify the power of governments' policy styles to explain policy failures; after comparison they had to conclude that: «The idea of policy style is intuitively appealing, and appears to capture a great deal of what we know about policy making in a range of political systems. Despite the face validity of the concept, [...] it is not effective in differentiating among countries. [...] These countries do at times display some of the behavior expected of them, but they are almost as likely to act in the opposite manner» (*ibid.*:647). Yet, it is consistent with a commitment to in-depth strategies that failure studies can conclude by claiming to have «revealed a good deal more about success and failure in governing» (*ibid.*:658), whereas having to acknowledge their inability to ascribe causality clearly.

The common understanding of the three tenets – policy failure as a linear yet multidimensional opposite of success; focus on remote governmental antecedents; necessarily overdetermined causality – imply an ontology and a methodology that make clear explanations hard. For clear causal analysis to become possible, hence, a shift is required. This is exactly what a Realist reframing can contribute to.

## **2. The heuristic gains from a Realist approach to failures**

In their seminal work, Bovens and 't Hart identify three common «implicit philosophies of governance» (1996:95ff) for making sense out of policy failures: Optimistic, Pessimistic and Realistic. Optimists see failings as anomalies of a working system of well-meaning actors, which can be avoided at the management level simply by providing the agents with better information. To the opposite, Pessimists believe failures to be inevitable effects of a debased political system – the solution to which, if any, can only result from

a radical change of the whole structures and working principles. Between these two poles, Realists maintain a third and intermediate stance that conceives «public policymaking and governance as essentially fragile activities» (1996:97) – i.e. complex processes where policy goals and tools are always disputed, shaped by «essentially self-centered, opportunistic, yet often mindless utility maximizers» (*ibid.*:98), and fuelled by competing values and ideologies. In such world, failures become normal occurrences: when ideologies result in interventions at odds with the good to produce, and when constraints prove unable to hold policy-shapers' opportunism effectively in check. Failures are hence intrinsic to real democratic dynamics, and mendable by proper changes in constraints.

These three stances provide observers with the cognitive core around which analyses revolve, so that to marry at least one of them is almost unavoidable. And indeed they all have currency in the public as well as in the scholarly discourse. From the perspective of explanatory research, however, the Realist alone seems fully equipped for casting an effective light on the causes of failures. Managers often neglect useful information because of reasons that stricter compliance may not remove; and radical changes can promise yet not guarantee that policy will succeed in a different future. Realism instead revolves around an unvarnished idea of the human nature, and ascribes policy failure to the relationship between policy ends and those behaviors allowed by the institutional constraints in use – and to some underlying «policy theory». The ontology of Realist governance so resonates with those scholarly positions that conceive policy design as the ground where the normative flavor of public interventions and their expected performance is decided (Schneider and Ingram 1990; Salamon 2002; Howlett 2002, 2011). To adopt it thus means to look for the efficient causes of policy performance in policy design first – rather than in management, constitutional design, or government. The consequences of such adoption are especially useful to the ontological shift from the consolidated failure framework.

Indeed, the Realist governance calls for the disentanglement of the usual analytical dimensions, as it recognizes that each dimension addresses a separate research question: the program, about the effectiveness of substantive and procedural tool-mixes in use; the procedural, about the legitimacy and consensus produced by the special shape of the decision-making processes for instrument choice; and the political, about the reasons that drove the relevant decision-makers to replace some tool or to keep it up. Indeed, these questions are theoretically nested (Howlett 2009); however, they generate distinct assessments against separate yardsticks. Such Realist disentanglement implies that the causal sequence assumed by the usual failure typology

– from the governmental antecedent to the constraints experienced by the policy-takers, to the outcomes and their politicization – is far from given. Rather, each connection between effectiveness, legitimacy and preferences has to be empirically proven against a proper outcome. As such, the Realist governance raises the question of how to prove such causal connection, so recalling a more general Realist ontology – the one set off by Bhaksar’s theory of science.

According to Bhaksar (2008), scientific ascription of causality cannot simply follow the detection of an empirical regularity – as perfect regularities hardly exist, and moreover not all of them are causal. Scientific ascription of causality cannot follow the simple application of a theoretical model to cherry-picked facts, either, as this strategy runs at the risk of some self-confirming selection of evidences. Nevertheless, causality can be grasped «scientifically»: by verifying an explanatory hypothesis about a generative mechanism.

This first requires a conjecture about the power of some property to make a difference to some other. The conjecture has to be deduced from the nature of the particular constituents that display the property; however, such nature is not investigated in itself: the structure of the particular constituents as well as the details of the generative process are both considered unobservable instead – which consents the analysis to «bottom out» and avoid the paradox of infinite regress (Glennan 2011). Although theory-driven, hypotheses about generative relations are hence little more than «educated guesses» in need of validation (Bhaksar 2008; Pawson 2004). Validation is nevertheless possible: at the metaphysical level, properties are true and universal – thus meaning that they allow for the «sameness» of the particular constituents which display them, although exclusively with respect to that property (Tooley 1987). In other words, different particulars can be expected to really behave the same, to the extent that they are characterized by the same degree of the same property, because of some underlying mechanism.

Having imagined a metaphysically sound causal hypothesis, then, does not end the scientific venture. Its heuristic capacity has to be confirmed – by developing a consistent explanatory model to test. This can prove far from easy, as to Realists mechanisms are necessary at some deep potential level, while their effects are not. So, explanatory models have to account not only for those cases where causality holds as expected, but also for those where the outcome is expected yet not displayed – because the proper antecedent is missing, or some other mechanism has defused the generative relation – as well as for cases where the outcome is displayed against expectations –



because the model has either ignored key causal properties, or the fact that they can have more than one empirical expression. To Realists, hence, actual explanations are really «complex»: conditional on antecedents, configurative in the sense that they depend on the joint occurrence of special properties, and disjunctive as different configurations can equally obtain.

Given the complexity of actual causation, Realist 'metaphysics for science' reasons that neater explanatory models can only result from a focus on special generative mechanisms and on their strictly related results. As such, it further justifies the disentanglement of research questions necessitated by the Realist ontology of governance. And indeed the separation of accounts of effectiveness from those of legitimacy and preferences entails further analytical gains from an explanatory perspective.

First, a focus on the specific mechanisms activated by policy design avoids the government-centered bias: a tool-based explanatory model of performance does not question the government's willingness, but the effectiveness of instruments alone. What matters here is whether the tools in use – regardless of the reasons beneath their choice – are enough to account for the delivered level of some «policy good» that political systems can be expected to provide (Almond and Powell 1982). Such an explanatory model is 'metaphysically' justified as all policy tools are capable of changing the utility function of the policy-takers, hence people's courses of action and, at the aggregate level, the policy outcome. However, each carrot, stick and sermon activates this change differently – by rewarding, punishing or persuading mechanisms. Each mechanism can so be supposed to obtain; at the same time, each mechanism embodies special beliefs about the degrees of freedom that policy takers shall be left to when coordinating for the production of that special policy goods – which affects its effectiveness (Eliadis *et al.* 2005; Salamon 2002). So, differences in the tool-mix in use can account for variations in an outcome, but also expose the actual results of previous games among competing values, and their stratification. The policy paradigms that may have had currency over time in a policy field can hence constitute a valid starting point for selecting the *explanans*.

Besides, Realist explanatory models require failures to be operationalized as neat «results», according to criteria external to the analytical model itself – i.e., as the expected outcome of a clear-cut regime rather than as the policy regime itself. In this way, a policy can be conceived as a way to organize the production of a policy good (Williamson 1996) that proper measures can gauge. So, the reification of policies as

given programs can be avoided, and the value of *explanans* restored to dimensions otherwise used for constructing the *explanandum*. The need for a clear outcome also indicates that explanatory analysis is not suited for uncovering the whole of the effects generated by a program: what can be 'realistically' tested are hypotheses about special generative mechanisms, i.e., about deep causes-effect relations, in an «analytical closure» necessary for causality to be clearly accounted for. This closure however does not exclude that outcomes can be built to render trade-offs among policy goals instead of single goals. Independent operationalization also allows the outcome to vary across cases – consistently with the assumption of 'Realist governance' that failures are normal results of running policies, yet not the only possible ones. The bias in the research design toward negative instances alone can so be corrected.

The resulting population of positive and negative cases then makes systematic comparison possible of the explanatory conditions, in line with the experimental rationale. The only problems left at this point are related to the technique for testing the model – which brings us back to the metaphysics of science.

### **3. The Qualitative Comparative Analysis as a method for Realist explanations**

Metaphysical Realism is especially demanding of explanatory models, as it requires them to account for any possible relationship between empirics and expectations. In so doing, Realism also sets a difference between 'ascertaining causality' and 'explaining' some outcome (Bhaksar 2008): the former operation demands that empirical evidence is provided of the deep law-like relation between some properties, while the latter entails consistent justifications of the ways in which such causality actually unfolded (von Wright 2009).

Within this perspective, to ascertain a causal relationship implies that a generative mechanism is understood as the potential, or the tendency, that some property has to bring about some outcome (Bhaksar 2008:74). It can hence be done by focusing on the regularities in the behavior of properties as variables, so that the net effect can be singled out of the independent on the occurrence of the dependent – or its average probability, been the effect of confounders discounted. So, causation can correspond to «Granger causation», and causes to statistical predictors – although not the reverse (Freese and Kevern 2013). Causation is thus framed in statistical terms as the increase or decrease in the average likeliness of an outcome. This means that a proven law-like relation tolerates outliers, i.e. cases where it does not obtain because

the underlying mechanism has somehow remained potential. As far as to account for such outliers exceeds its heuristic ambition, probabilistic analysis best fits the goal of gauging some properties' potential beneath actual occurrences, on the basis of empirical evidence but at the same time beyond it. Its scientificity is guaranteed until a proper design and an accurate determination of the model are provided that make the operation as consistent to the experimental rationale as possible, and that minimize the problem of omitted variables (Rubin 2005).

Metaphysical Realists also maintain that, differently from causation, explanation has to account for each and every actual occurrences of the *explanandum*. As such it requires a model and a strategy able to cover both the conditions under which the law-like relation actualizes, and those under which the relation fails to trigger. Such model and strategies hence have to safeguard «the hold on the individual case» (Scriven 1959:467) which a statistical understanding does not allow: probabilistic laws cannot describe the actual patterns of single events because – bluntly – each event is not determined before it is caused to happen by a contextual balance of forces (Bhaksar 2008:95ff.). Realist explanations account for «dead causes» alone: thus cannot be but deterministic, ex-post, and valid within the boundaries of a given population of cases. Also, they can take the form of conjunctures of conditionals, thus generating as many distinct configurations as required to account for special subgroups of cases with the same outcome – even for a single case. The identification of such configurations can nevertheless be scientific as far as it relies on Mill's recognized canons for proving that the properties truly are conditional – material, or counterfactual. Provided again a proper research design, cross-case systematic comparisons can then be used to verify that the properties in the model are necessary or sufficient conditions to the occurrence of the outcome. And indeed the recognition of necessity and sufficiency is a logical operation, rather than probabilistic (Tooley 1987). As such, it can be treated in set-theoretical terms. Necessary conditions are those without which an outcome cannot occur, but that do not guarantee that it will: hence, given a population, a condition is necessary when the cases that display it are a superset of the cases displaying the outcome. Similarly, sufficient conditions are those which are always associated to the occurrence of the outcome, yet without which the outcome can still occur (because of some other triggering condition): so, a condition is sufficient when the cases that display it are a subset of the cases displaying the outcome (Ragin 1987).

As far as Realist explanations are configurations of necessary and sufficient conditions accounting for the actual occurrence of an outcome within a population, the set-theoretical treatment of Qualitative Comparative Analysis (thereafter, QCA: Ragin

1987, 2008; Schneider and Wagemann 2012) appears especially suited for uncovering them.

QCA is neither a case-oriented nor a variable-oriented research strategy: in his inventor's words, is rather «diversity-oriented» (Ragin 2000). It aims to explain the distribution of an outcome within a population by configurations of necessary and sufficient conditions deemed causal after some theory – in Realist metaphysical terms, after some hypothesis about the triggering of special generative mechanisms; in Realist governance terms, after these policy paradigms that over time may have framed the policy good problem and tipped the balance in favor of special tools because of their productive capacity. The appropriateness of treating such Realist hypotheses by QCA in policy performance studies is underlined by Ragin himself, to whom «central to policy is determining which causal conditions are decisive in which contexts» (2006:18) – i.e., generating «circumstantially delimited» accounts of «the conditions that enable or disable specific connections between causes and outcomes» (ibid.:17).

The starting point of QCA hence is an array of conditions mechanistically relevant to the occurrence of an outcome, which can basically be either present (1) or absent (0). A truth table is then derived that «exhausts» the possible combinations of presence and absence of each condition – i.e., the logically possible types. Cases from a meaningful population are then assigned to the related type in the truth table depending on the configuration displayed; the type is so associated an actual outcome. Whatever the expectations, the generative power is so revealed empirically – and not of a single condition, but of a whole configuration as instantiated in actual cases. These generative configurations can also include the absence of some property – which brings Mill's «preventing or counteracting causes» into the explanatory picture, too. Moreover, the types with the same outcome are thus all proven 'productive', so that each constitutes an «equifinal», «primitive Boolean» causal path.

Primitive configurations can however be redundantly detailed. Pairwise comparisons then «minimize» equifinal primitive paths, by dropping these conditions alone which demonstrate irrelevant unless a minimal number is reached of «prime implicants» that cannot be further reduced and that cover as many of the primitive configurations as possible (Ragin 1987:93ff). Prime implicants usually then consist of distinct and more general configurations, telling which essential combinations of conditions from the original model are enough to explain which subpopulations. Prime implicants hence are local «recipes» for the outcome – determinist, configurational, and valid within the boundaries of the population under scrutiny at the level of the single

case, consistently with the requisite of metaphysical Realism for good explanations. Moreover, the set-theoretical and Boolean treatment simply establishes that the essential implicants in the solution are of necessary and sufficient conditions to lead to the outcome, thus confirming the initial generative hypothesis: but it does not exclude that other mechanisms insist on the outcome, hence that other explanations may be possible. However, given the analytical closure, such possibility is simply irrelevant: to the extent that the model in use accounts for each and every case in the population, the generative hypothesis is verified. On the contrary, a single unaccounted case is enough to indicate that the model is underdetermined, and a single contradictory case enough to confute the model – first its operationalization, then the hypothesis about the generative conditions, and eventually the basic theory about the nature of the unobserved constituents (Rihoux and de Meur 2008).

#### **4. An example: explaining the inability to deliver green growth in the EU15**

Metaphysical Realist explanatory models imply a comparative research design where the explanans include all the generative conditions possibly leading to an outcome after a special mechanism. Governance Realist explanatory models narrow the generative conditions down to the properties of policy tools as far as they can reshape the policy-takers' utility function and so their policy-relevant behaviors. The tools and the properties which deserve attention cannot however be selected randomly. To enter the explanatory model, the assumption of their relevance has to be backed. Policy ideas do offer such backing, as policy theories defines the appropriateness of special tools to a special policy good that can push a policy-maker to introduce them in the field (Blyth 2002). Realist policy models can hence develop around the many policy ideas which may have had currency over time – around their framing of the problem and the related instruments.

In the field of green policy, ideas mainly revolve around the viability of instruments because of their relationship with economic growth. In the 1960s, when green concerns first entered the agenda, standard economics stated there was a trade-off between the two goals: social gains could only come at the price of private losses, which meant that governments could only choose between green pauperization and dirty wealth. The theory was deemed so relevant and convincing that, as a result, interests and the public were assumed to polarize around a new social cleavage pitting modern 'productionists' against post-modern 'environmentalists' (Inglehart 1990). In the 1990s,

Michael Porter challenged this theory at its very roots. In his view, pollution is a loss in market efficiency, and green policy may instead result in joint social and private benefits if constraints are designed so as to induce policy-takers to behave properly (Porter and van der Linde 1995). Despite criticisms and far from conclusive evidence, Porter's theory has been widely adopted, often under the label of «sustainable development», especially at the international and European level (Lenschow and Zito 1998; COM(2010) 2020 fin, p.5). From these stances Colby (1991) recognizes five possible special balances of growth and environmental protection and related tool-kits which had currency over time, namely:

- «Frontier economics» (FE), which revolves around the baseline axiom of unlimited or substitutable natural resources, and justifies their extensive exploitation for the sake of growth and employment. Productionist beliefs dismiss environmental concerns so that no specific green policy is envisaged. The hallmark is national economic market-friendly regulation only, which is expected to result in a joint rise in resource depletion and economic growth.
- «Deep ecology» (DE), which develops the position environmental movements took in reaction to Frontier Economics. It reverses the productionist core belief by assuming the environment to be the only relevant goal. Growth is deemed to be the problem in need of restraint for the sake of conservation. The main tool is command-and-control regulation with the deliberate purpose of community protection and de-growth. The paradigm predicts high environmental results with negative economic performances.
- «Environmental protection» (EP), which characterizes those early green policies that simply correct FE at the margins. Growth is still the main priority, but environmental harm enters the policy agenda in the form of market failures calling for government interventions. The tools are command-and-control environmental regulation, and expenditures for ex-post 'medical' treatment – the burden of which lies with the national taxpayer collective via universal taxation. The expected results are some depletion constraints but also some reduction in growth.
- «Resource management» (RM), which involves a true ideational shift because it believes green growth possible. In line with Porter's theory, environmental harm is seen as externalities that can be handled within markets, by defining the individual responsibility for depletion and by fostering innovation. Thus, in a context of market-friendly regulation, technological change is made a reasonable strategy by both the stick of 'polluters pay' taxation and the carrot of targeted spending. The expectation is that depletion decreases while growth increases.

- «Eco-development» (ED), which goes further by believing that the environmental rationale is better able to shape economic mechanisms. The basic idea is that 'preventing pollution pays', and that this becomes possible as social processes are re-designed to function as ecosystems. The instrumentation replicates the Resource Management toolkit; however, the perspective moves from individual to collective responsibility, and community boundaries become more relevant. The expectation is again for a decoupling of growth and pollution trends, although the results would be less remarkable.

If differences among policy theories refer to the trade-off between environmental and economic policy goals, then a proper measure of performance shall not focus on pollution alone. Rather, it can gauge whether and to what degree the two productions – of growth as «economic goods» and resource depletion as «environmental bads» – have «decoupled» (OECD 2002).

Differences in such outcome may then be ascribed to differences in the incentives and restraints set by special tool-mixes (Salamon 2002). In them, what matter is whether Regulation is (1) coercive or soft; (2) pro- or anti-competition (3) environmentally protective; Taxation is targeted to polluters; and Spending aims to R&D innovation (Tab1).

**Tab 1.** Paradigms by policy priority, distinctive attributes and expected outcomes.

		<b>Frontier Economics</b>	<b>Environmental Protection</b>	<b>Resource Management</b>	<b>Eco-Development</b>	<b>Deep Ecology</b>
<b>Priority:</b>	<i>economic</i>	high	high	high	moderate	none
	<i>environmental</i>	none	moderate	high	high	high
<b>Attributes:</b>	<i>soft</i>	present	absent	present	present	absent
	<i>competitive</i>	present	present	present	absent	absent
	<i>protective</i>	absent	absent	present	present	present
	<i>targeting</i>	absent	absent	present	absent	absent
	<i>innovating</i>	absent	absent	present	present	absent
	<i>evaluating</i>	enhancing the rationale of the tool-mix when present				
<b>Outcome:</b>	<i>decoupling</i>	no	no	yes	yes	no

Source: Author's operationalisation of Colby (1991).

All these substantive tools can however prove ineffective unless some procedural device is laid down to keep them accountable and on track (Howlett 2002). Ex-ante policy analysis may help regulators set up 'sticks' appropriately and learn about the ways

they work in the field (Radaelli 2009), while performance targets have positive effects on ‘carrots’ (Poterba and von Hagen 1999). Together, they are expected to magnify the rationale of the tool-mix in use and secure the expected outcome.

In the EU15 – considered as a relatively homogeneous population of cases because of the pressures exerted by the European environmental directives yet different enough in their environmental approach to allow for meaningful comparisons – the above selection of *explanandum* and *explanantes* generate the operationalization in tab.2, as explained in the related notes.

**Tab.2.** Policy tools and green growth: raw and crisp values, selected countries.

	<i>gauge</i>	<i>outcome</i>	<i>regulation</i>			<i>taxation</i>	<i>spending</i>	<i>governance</i>		
		<i>decoupling</i>	<i>soft</i>	<i>competitive</i>	<i>protective</i>	<i>targeting</i>	<i>innovating</i>	<i>steering</i>	<i>assessing</i>	<i>evaluating</i>
<b>AUT</b>	<i>raw</i>	0.021	2.338	0.861	11.000	1.597	1.533	0.475	4.300	
	<i>crisp</i>	1	0	0	1	0	0	1	0	0
<b>BEL</b>	<i>raw</i>	0.026	4.471	0.309	10.000	8.192	2.133	0.091	1.000	
	<i>crisp</i>	1	0	1	0	1	1	0	0	0
<b>DNK</b>	<i>raw</i>	-0.024	1.267	0.788	7.000	17.377	2.033	0.459	8.000	
	<i>crisp</i>	0	1	0	0	1	0	1	1	1
<b>FIN</b>	<i>raw</i>	0.034	1.292	0.455	13.000	1.471	2.033	0.619	9.300	
	<i>crisp</i>	1	1	1	1	0	0	1	1	1
<b>FRA</b>	<i>raw</i>	0.032	2.900	0.555	7.000	5.330	3.000	0.377	4.000	
	<i>crisp</i>	1	0	0	0	1	1	0	0	0
<b>DEU</b>	<i>raw</i>	0.028	1.725	0.827	7.000	0.000	3.300	0.231	4.700	
	<i>crisp</i>	1	1	0	0	0	1	0	0	0
<b>IRL</b>	<i>raw</i>	0.013	3.725	0.344	10.500	0.142	1.733	0.450	7.300	
	<i>crisp</i>	0	0	1	0	0	0	1	1	1
<b>ITA</b>	<i>raw</i>	0.016	1.825	0.736	14.500	1.062	2.500	0.362	3.300	
	<i>crisp</i>	0	1	0	1	0	1	0	0	0
<b>NLD</b>	<i>raw</i>	0.040	1.483	0.443	10.000	17.374	2.567	0.575	8.300	
	<i>crisp</i>	1	1	1	0	1	1	1	1	1
<b>PRT</b>	<i>raw</i>	-0.017	1.772	0.561	17.000	0.000	3.500	0.241	2.300	
	<i>crisp</i>	0	1	0	1	0	1	0	0	0
<b>ESP</b>	<i>raw</i>	0.015	4.408	0.616	23.000	0.977	2.133	0.160	1.000	
	<i>crisp</i>	0	0	0	1	0	1	0	0	0
<b>SWE</b>	<i>raw</i>	0.038	2.183	0.560	14.000	2.349	1.433	0.544	1.000	
	<i>crisp</i>	1	0	0	1	1	0	1	0	0
<b>GBR</b>	<i>raw</i>	0.036	1.842	0.241	7.000	2.881	1.733	0.554	9.700	
	<i>crisp</i>	1	1	1	0	1	0	1	1	1

**Notes:** of the EU15 Member countries, Greece is dropped because of unreliable data; and Luxembourg because of its peculiar dimensions and policy performance which make it a little comparable case.

As for conditions:

- *Decoupling* is based on two different gauges of depletion related to the Gross Domestic Product (GDP) – all taken from Eurostat (2011). GDP is measured as chain-linked volumes at 2000 exchange rates, which allows the within-country comparison implied by the calculation. The first measure of depletion is the national annual production of greenhouse gas emissions (GGE), usually used to assess countries’ progress toward the Kyoto targets and deemed robust enough for comparison (OECD 2002). Yet GGE alone is not enough to appreciate the greening of economic processes on which the paradigms of Resource Management and Eco-



Development are based: hence the need for a second decoupling factor, measured by Domestic Material Consumption (DMC). DMC quantifies the annual amount of raw materials used by the national economy (except water and air, for which data are not robust enough), thus indicating how environmentally efficient national production and consumption are. The factors are each calculated on a yearly base for the period 2004-2006. Their averages are then equally weighted into a synthetic 'decoupling' index so as to gauge the overall national system performance.

- *Soft* is a measure from the OECD (2011), namely a sub-index of its wider indicator of Product Market Regulation. Its value refers to the use of command-and-control regulation in 2003 in the whole economy, not simply to the environmental sector – since the regulatory style is unlikely to be sectoral (Salamon 2002).

- *Competitive*. The degree to which the rules in use hinder economic competition is quantified by a different sub-index of the OECD (2011), based on the barriers that overall domestic regulation exerted on trade and investments in 2003.

*Protective*. The governments' actual commitment to conservation is based on average 2003-2004 data for the protected terrestrial areas for biodiversity as a percentage of the whole country extension (Eurostat 2011).

- *Targeting*. Since the relevant analytic point lies in the extent to which the tax base targets pollution, the measure is calculated as the average ratio in 2002-2004 between targeted revenues and the total green fiscal levy, using Eurostat (2011) data.

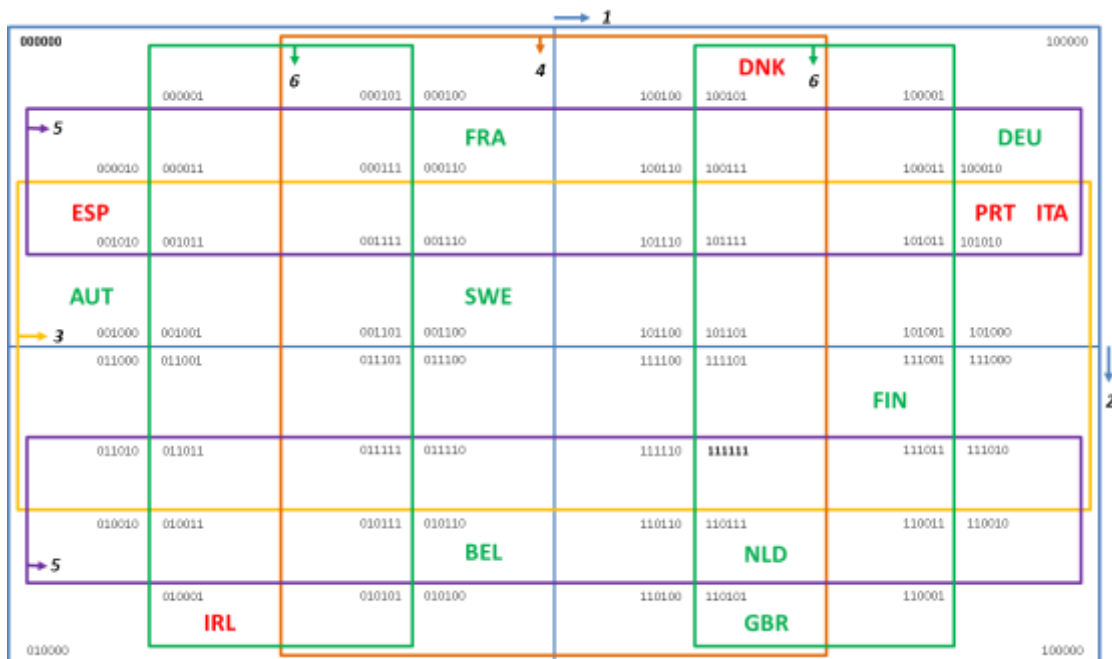
- *Innovating*. The condition enters the model as the 2002-2004 average ratio of environmental Resource and Development appropriations out of the total green outlay, based on Eurostat data. These data only allow meaningful comparisons to be made for public spending levels, leaving private investments out of the picture. However, this choice is in line with the research focus on governments' strategies alone.

- *Evaluating*. The quantity is based on: the Bertelsmann Stiftung's indicator of the evidence-based instruments in use in 2008 (Bertelsmann Stiftung 2009); and on the use of performance budgeting systems by the central governments in 2005, as calculated by the OECD Performance Budgeting Procedure Survey (OECD 2007).

To turn «raw measures» into «crisp scores», crossover threshold is set after the «natural gaps» identified by Tosmana as follows: *decoupling*: 0.02; *soft*: 2; *competitive*: 0.5; *protective*: 10.9; *targeting*: 2; *innovating*: 2.1; *steering*: 0.4; *assessing*: 5. *Evaluating* results from the intersection, hence the Boolean conjunction, of *steering* and *assessing* – thus it takes the minimum value.

Each case hence makes «empirically true» one among the  $2^n$  possible combination resulting from  $n$  conditions, as in Fig.1.

**Fig.1.** Venn diagram of policy tools for green growth



**Keys:** Red cases display the negative outcome, green cases the positive.

- 1= *soft*, present (1) on the left and absent (0) on the right;
- 2= *competitive*, present (1) in the top half, absent (0) in the bottom;
- 3= *protective*, present (1) in the horizontal large rectangle, absent (0) outside;
- 4=*targeting*, present (1) in the vertical large rectangle, absent (0) outside
- 5=*innovating*, present (1) in the horizontal small rectangles, absent (0) outside
- 6=*evaluating*, present (1) in the vertical small rectangles, absent (0) outside

The same figure also reveals that the cases of the population are not enough for providing each type in the truth table with an empirical outcome. These empty types or «logical remainders» may however be used as counterfactual conditional to minimize the empirically true types (Ragin 1987, 2008). When such counterfactuals are minimized after «directional expectations» about the contribution of each condition to the outcome, the model generates the following «intermediate solutions» for policy failures of green growth policies:

- innovating•protective•competitive → decoupling (DNK)
- innovating•protective•soft•targeting → decoupling (IRL)
- INNOVATING•PROTECTIVE•competitive•targeting•evaluating→ decoupling (ITA, PRT, ESP)

These «recipes» demonstrate far from symmetric with respect to the intermediate solutions for policy success:

- innovating•PROTECTIVE → DECOUPLING (AUT, FIN, SWE)
- INNOVATING•protective•TARGETING → DECOUPLING (BEL, FRA)
- INNOVATING•protective•SOFT: → DECOUPLING (DEU)
- EVALUATING•TARGETING•COMPETITIVE•SOFT → DECOUPLING (NLD, GBR)

The solutions demonstrate that the policy mixes with explanatory power do not precisely overlap any pure environmental paradigm; moreover, in revealing the actual triggering and enabling mechanisms at work in each case, they contribute to refine expectations about instruments and mixes.

On the side of failure, solutions highlights how, regardless of the state of the economy, the absence of green constraints is actually associated to not decoupling (IRL, DNK); however, the joint presence of protective regulation and expenditure for innovation does not deliver as expected when evaluation is missing (and the economy is not competitive, and green taxation is designed as a collective responsibility: ITA, PRT, ESP). On the opposite side of success, decoupling can occur after, alternatively, protective regulation without public spending (AUT FIN SWE); or non-protective

regulation with public spending and either targeted taxation (BEL, FRA) or soft economic regulation (DEU); or the 'neo-liberal regulatory' part of the resource management paradigm, when subjected to evaluation (NLD; GBR).

Results also make two more general points. First: policy failures deserve a separate analysis as they cannot be explained by the opposite solutions to success. Second: as all the cases are covered by some causal path, that substantive and accountability policy tools are enough for a valid explanatory model of policy failures as poor delivery. Consensus and legitimacy issues hence can be left out of this step of the analysis.

## 5. Conclusions

Because of its special ontological and methodological assumptions, the consolidated framework for policy failures shows some limit when used for explanatory purposes. For explanation to become possible, the article has reasoned that a Realist shift is required. Such shift demands that the three dimensions in framework – program, procedure and politics – are first disentangled around three distinct research questions – about, respectively, effectiveness, legitimacy and consensus – and then reorganized as *explanans* of clear-cut *explananda*.

Explanation is also a different operation from proving causality: despite both can stem from a hypothesis about some basic generative mechanism, the latter aims to establish deep law-like cause-effect relations beneath the vagaries of empirical regularities and hence is consistent with a probabilistic treatment, while the former precisely aspires to account for these vagaries. Explanations thus imply an inquiry into those essential conditions which actually trigger some deep mechanism, or defuse it; as such, they cannot be but ex-post exercises within a given population. The identification of the essential triggering and defusing conditions then requires a technique that still secures scientific results while treating evidences properly – i.e. logically.

The QCA is demonstrated such suitable technique, as it not only starts from the assumptions that the occurrence of an outcome depends on configurations, but also can identify which of them are the essential conditions necessary and sufficient for an outcome to actually occur, consistently with a quasi-experimental design.

Although the demonstration is run only on one of the three dimensions in the consolidated framework, it is easy to imagine the applicability of QCA to the other

research questions, too – once that a proper outcome is provided, and the underlying mechanisms identified and operationalized as a set of relevant condition. That is why the Realist Configurative approach is deemed a fruitful direction for future failure studies.

## References

- Almond, G.A. and Powell, G.B. (1982), 'Evaluating political goods and productivity', *International Political Science Review*, 3(2): 173-181.
- Bhaskar, R. (2008 [1974]), *A Realist Theory of Science*, Oxon: Routledge.
- Blyth, M. (2002), *Great Transformations: Economic Ideas and Institutional Change in the Twentieth Century*. Cambridge: Cambridge UP.
- Bovens, M.A.P., 't Hart, P. and B.G. Peters (1998), 'Explaining policy disasters in Europe: Comparisons and reflections', in Gray, P. and 't Hart, P. (eds.), *Public Policy Disasters in Western Europe*, London: Routledge, 199-220.
- Bovens, M.A.P., 't Hart, P. and B.G. Peters, (eds.) (2001), *Success and Failure in Public Governance*, Cheltenham: Edward Elgar.
- Bovens, M.A.P., 't Hart, P. (1996), *Understanding Policy Fiascoes*, New Brunswick: Transaction.
- Carlsson, L. (2000), 'Policy Networks as Collective Action', *Policy Studies Journal*, 28(3): 502-520.
- Colby, M.E. (1991), 'Environmental management in development', *Ecological Economics*, 3 (3), 193-213.
- Dunleavy, P. (1995), 'Policy disasters: explaining the UK's record', *Public Policy and Administration*, 10(2): 59-64.
- Eliadis, P., Hill, M.M., and M. Howlett (2005), *Designing Government*, Montreal: McGill-Queen's Press.
- Elman, C. (2005), 'Explanatory typologies in qualitative studies of international politics', *International Organization* 59(2): 293-326.
- Eurostat (2011), *Database*, [http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search\\_database](http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database).
- Freese, J and Kevern, J.A. (2013), 'Types of causes', in Morgan, S.L. (ed), *Handbook of Causal Analysis for Social Research*, New York: Springer, 27-41.
- Fung, A, (2006), 'Democratizing the policy process', in Moran, M., Rein, M., and Goodin, R. E. (eds.), *The Oxford Handbook of Public Policy*, Oxford: Oxford UP, 669-687.
- Glennan, S. (2011), 'Singular and general causal relations: A mechanist perspective', in Illari, P. M., Russo, F., and Williamson, J. (eds.), *Causality in the Sciences*, Oxford: Oxford UP, 789-817.
- Gray, P. (1996), 'Disastrous explanations or explanations of disaster: a reply to Patrick Dunleavy', *Public Policy and Administration*, 11(1): 74-82.
- Gray, P. (1998), 'Policy disasters in Europe: An introduction', in Gray, P. and 't Hart, P. (eds.), *Public Policy Disasters n Western Europe*, London: Routledge, 3-20.
- Howlett, M. (2000), 'Managing the "hollow state": Procedural policy instruments and modern governance', *Canadian Public Administration*, 43(4), 412-431.
- Howlett, M. (2009), 'Governance modes, policy regimes and operational plans: A multi-level nested model of policy instrument choice and policy design', *Policy Sciences*, 42(1), 73-89.
- Howlett, M. (2011), *Designing public policies: principles and instruments*, New York: Routledge.
- Inglehart, R. (1990), *Culture shift in advanced industrial society*, Princeton: University Press.
- King, A. (1975), 'Overload: problems of governing in the 1970s', *Political Studies* 23:(2-3):284-296.
- Lundahl, M. and M.L. Wyzan (2005), 'Introduction', in *Id.* (eds), *The Political Economy of Reform Failure*, London: Routledge, 1-22.
- Matland, R.E. (1995), 'Synthesizing the Implementation Literature: The Ambiguity-Conflict Model of Policy Implementation', *Journal of Public Administration Research and Theory*, 5(2): 145-174.
- McConnell, A. (2010a), 'Policy success, policy failure and grey areas in-between', *Journal of Public Policy*, 30(3): 345-62.

- McConnell, A. (2010b), *Understanding Policy Success*, Basingstoke: Palgrave Macmillan.
- Moe, T.M. (1995), 'The politics of structural choice: Toward a theory of public bureaucracy', in Williamson, O.E. (ed.), *Organization Theory*, Oxford: Oxford UP, 116-153.
- Moran, M. (2001), 'Not steering but drowning: Policy catastrophes and the regulatory state', *The Political Quarterly*, 72(4): 414-27.
- Most, B.A. & Starr, H. (2003), 'Basic logic and research design: Conceptualization, case selection, and the form of relationship', in Goertz, G. & H.Starr (eds), *Necessary Conditions*, Lanham: Rowman & Littlefield, 25-46.
- Niskanen, W.A. (1971), *Bureaucracy and Representative Government*, New York: Aldine Atherton.
- OECD, (2002), Sustainable development, SG/SD(2002)1/final.
- OECD, (2007), *International database of budget practices and procedures*, [www.oecd.org/gov/budget/database](http://www.oecd.org/gov/budget/database) .
- OECD, (2011), *Product market regulation database*, [www.oecd.org/economy/pmr](http://www.oecd.org/economy/pmr) .
- Olson, M. (1982), *The Rise and Decline of Nations*, New Haven, Yale University Press.
- Pawson, R. (2004 [1989]). *Measure for Measure: Manifesto*, London: Routledge.
- Porter, M.E. and van der Linde, C. (1995), 'Toward a new conception of the environment-competitiveness relationship', *Journal of Economic Perspectives*, 9 (4), 97-118.
- Poterba, J. M. and von Hagen, J. (eds.) (1999), *Fiscal institutions and fiscal performance*, Chicago: University Press.
- Radaelli, C.M. (2009), 'Measuring policy learning', *Journal of European Public Policy*, 16 (8), 1145-1164.
- Ragin, C.C. (1987), *The comparative method*. Berkeley: University of California Press.
- Ragin, C. C. (2000), *Fuzzy-set social science*. University of Chicago Press.
- Ragin, C.C. (2006), The limitations of net-effect thinking. In: B. Rihoux and H. Grimm, eds. *Innovative comparative methods for policy analysis*. New York: Springer, 13-42.
- Ragin, C.C. (2008), *Redesigning social inquiry*. Chicago: University Press.
- Rihoux, B. and de Meur, G. (2008), 'Crisp-Set Qualitative Comparative Analysis', in Rihoux, B. and Ragin, C.C. (eds.), *Configurational Comparative Methods*, London, Sage, 33-68.
- Rubin, D. (2005), "Causal Inference Using Potential Outcomes", *Journal of the American Statistical Association*, 100 (469): 322-331.
- Salamon, L.M. (ed.), 2002, *The Tools of Government*, Oxford: Oxford University Press.
- Schneider, A. and Ingram, H., 1990. Behavioral assumptions of policy tools. *The Journal of Politics*. 52(02): 510-529.
- Schneider, C.Q. and Wagemann, C. (2012), *Set-Theoretic Methods for the Social Sciences*, Cambridge: Cambridge UP.
- Scriven, M. (1959), 'Truisms as the Grounds for Historical Explanation', in Gardiner, P. (ed.), *Theories of History*, Glencoe: The Free Press, 443-75.
- Tooley, M. (1987). *Causation: A Realist Approach*. Oxford: Oxford UP.
- Vining, A.R. & Weimer, D.L. (1990), 'Government supply and government production failure: A framework based on contestability', *Journal of Public Policy*, 10(1): 1-22.
- von Wright, G.H. (2009 [1971]), *Explanation and Understanding*, London, Routledge.
- Williamson, O.E. (1996), *The Mechanisms of Governance*, Oxford, Oxford UP.
- Wolf, C. (1987), Market and non-market failures: Comparison and assessment', *Journal of Public Policy*, 7(1): 43-70.