Disproportionate responses to European Union policy?

Customized problem-solving in the food safety sector

Dr. Eva Thomann, Institute of Political Science, Heidelberg University

eva.thomann@ipw.uni-heidelberg.de

Abstract

This paper assesses how the customization of EU food safety policies affects whether policies are successfully implemented into practice. Customization means that member states use their discretion to adapt EU rules during legal transposition. Policy implementation scholars and practitioners hold contradictory views on the role of discretion, framing it either as a control problem, or as a crucial precondition for effective problem-solving. Simultaneously, the practical implementation of EU policies remains largely a black box. This paper applies fuzzy-set Qualitative Comparative Analysis to compare the implementation of 19 EU veterinary drugs rules in four Western European member states and Switzerland (N= 95). Results show no straightforward relationship between customized policies on paper and the successful delivery of policy outputs and outcomes. Still, the complex configurations that facilitate successful implementation very often entail extensively customized EU policies. Conversely, an (almost) literal transposition of EU food safety policies is a prominent part of the story explaining unsuccessful implementation. This evidence encourages us to rethink the paradigm of avoiding "over-implementation" in the EU regulatory state.

Keywords: Customization, discretion, Europeanization, gold-plating, policy outcomes

Introduction

"The closer one is to the source of the problem, the greater is one's ability to influence it; and the problem-solving ability of complex systems depends not on hierarchical control but on maximizing discretion at the point where the problem is most immediate."

Richard F. Elmore (1979: 605)

Increased resistance to antibiotics represents a serious risk to human health. One major way in which antibiotic resistance emerges and travels to human consumers is when too many antibiotics are given to livestock. To address this problem in the European single market for food products, the European Union (EU) has taken several measures. For example, the EU restricts the use of antibiotics for animals to the amount needed for one treatment. However, as to 2011, different countries interpreted this rule in strikingly diverse ways. In Germany, the use of antibiotics is restricted to seven days; in Austria, to one month; and France and the UK have simply adopted the EU wording without specifying a time limit. Eating meat or eggs, for instance, might hence be more or less safe, depending on the interpretation of the EU rule in the country of origin. This article scrutinizes the results of such interpretations – labelled "customization" (Thomann 2015a) - and their effects on the domestic outcomes of EU policy.

The European Union (EU) as a regulatory state seeks to reconcile integration with legitimate differences in national preferences. In such a complex multilevel governance structure, decision-makers devise centralized policies as a response to joint policy challenges, for instance, to ensure food safety. However, recent research highlights that some member states "customize" EU policy when implementing it, by adding or reducing the amount and stringency of the respective rules. Throughout the implementation chain, EU rules undergo a process of vertical regulatory change that results in divergent domestic policies (Bauer and Knill 2014).

1

This "legitimate diversity" is a central aspect of the European experience (Héritier 1999; Majone 1999; Thomann and Sager 2017a, b). The concept of customization captures how countries use their discretion to interpret EU policy and adapt it to local circumstances during implementation, resulting in a diversity of tailor-made domestic solutions to shared policy problems.

This study confronts top-down and bottom-up implementation perspectives on the role of the use of discretion for explaining differences between countries in the practical application (outcomes) of European Union (EU) law. Top-down perspectives view discretionary adaptations as a potentially problematic transposition outcome that can conflict with compliance. From a top-down implementation perspective, the EU policy is the adequate solution to a given policy problem– deviations (customization) could imply red tape or distortions of competition that create unnecessary burdens for businesses ("gold-plating"). Correspondingly, earlier studies have termed the phenomenon "over-implementation". From such a view, customization indicates that member states overinvest in an EU policy. Conversely, bottom-up implementation theory emphasizes that member states' closeness to the source of the policy problem enhances their ability to improve EU policies, by adapting them to heterogeneous contexts and integrating local policy positions. A bottom-up view thinks of member states who customize EU policy as especially eager problem-solvers, going even beyond the EU in investing in a policy solution. From such a view, member states correct for inadequate policy responses at the EU level. Country-specific implementation can increase the likelihood of effective policy solutions at the point where the problem is most immediate.

In confronting these two contrasting view, this paper asks: how does the customization of EU food safety policy affect successful policy implementation? In order to determine whether and how customization has a beneficial or distorting role, EU and domestic policies are set in relation with practical policy outputs and outcomes. The role of customization for successful

policy implementation is embedded into an explanatory framework that lends from Sabatier and Mazmanian (1980) to account for problem tractability, the degree of centralization of the implementation structure, the enforcement strategy, the coherence of the policy design, and domestic opposition to EU rules. The empirical case at hand is the implementation of 19 EU food safety rules in four member states and the differentially integrated non-member state Switzerland (N = 95). The analysis is based on fuzzy-set Qualitative Comparative Analysis (QCA), complemented with targeted analyses of typical and deviant cases.

The paper proceeds as follows. A first section briefly introduces the customization concept. The second section tackles the question of how we can think of successful implementation in the EU, and how it can be measured given conceptual considerations as well as the data situation. Subsequently I derive an explanatory framework that draws from classical implementation theory and integrates salient insights of Europeanization research (Knill 2015; Treib 2014). I then discuss the data and methods. Section six presents the results. I conclude by discussing implications for successful policy implementation in the EU, and multilevel implementation theory more generally. In tracing the implementation of EU directives from transposition to domestic outputs and outcomes, the paper contributes significantly to understanding the practical solution of common problems in the EU's multi-level regulatory system.

Customization: discretion in multi-level policy implementation

Customization refers to discretionary adaptations of EU rules by member states during transposition. We can think of customization as vertical changes that EU policy undergoes during transposition (Adam et al. 2015; Schaffrin et al. 2015; Thomann and Zhelyazkova 2017). These changes occur regarding the amount of rules (density) and their content (intensity). Customization here is defined as the degrees to which the domestic regulations implementing an EU rule in a compliant manner complement the latter with *more or stricter rules* than

3

required. *Customized density* captures the amount/ number of rules (policies and/or instruments) that are added to the EU provision (e.g., scope restrictions, concretizations, requirements, exemptions or specifications added or taken away). A domestic transposing rule customizes the Directive provision in its density, if it adds at least one rule/standard/condition/specification/exemption to the latter. *Customized restrictiveness* captures how EU rules are adapted in the extent of the substantial freedom left to policy addressees, or the generosity of the services or resources provided (Bauer and Knill 2014: 33; Knill et al. 2012: 430).

To measure customization, I adopt Knill and Lehmkuhl's (2002) distinction of 'institutional compatibility' to measure the extent of *occurred* (as opposed to required) domestic changes in *policies* (as a subset of domestic arrangements) in response to EU policies. Each domestic regulation is classified according to whether changes in comparison to the EU policy are absent (0), *relatively* moderate (1) or extensive (2) concerning a) density and b) intensity. Moderate changes usually entail 1, and extensive, 2 or more changes. These two dimensions are added into a customization index. With values of 3 or 4, customization is fully extensive. Without any changes or only moderate changes on one dimension, customization is limited. Values of 2, representing either only moderate changes or no changes on one dimension and extensive changes on the other, indicate intermediate customization.

How does customization affect successful policy implementation?

Policy implementation theory assumes a strong link between the use of discretion by implementing actors, on the one hand, and the success implementation of the policy, on the other (see Matland 1995; Shapiro 1999). We find contradictory views on the exact role of discretion (Hupe 2013; Matland 1995; Rutz et al. 2015; Thomann et al. 2016; Winter 2003).

On the one hand, top-down perspectives tend to treat deviations from the policy-on-paper as a

control problem, making it increasingly likely that policy means and ends will be mismatched (Howlett 2004: 5). When policy implementers pursue their own goals rather than those of the policymaker, this can undermine the effectiveness and democratic legitimacy of a program (Tummers and Bekker 2014). Thus, from a top-down view, extensive customization should negatively affect implementation success. This resonates with the discussion surrounding the "gold-plating" or "over-implementation" of EU policies: when member states regulate more than minimally required by the EU, this creates unnecessary burdens for businesses and other target groups and is hence predominantly seen as problematic. However, the actual empirical consequences of gold-plating have only been assessed in a few case studies which did not identify a wide-spread problem (Davidson 2006; Falkner et al. 2005; Jans et al. 2009; Morris 2011; Versluis 2003, 2007; Voermans 2009).

Conversely, bottom-up perspectives have a very different view on discretion. They see policy implementers as de facto policy makers and problem-solvers who are responsive to the specific context and needs of target groups (Brodkin 2011; Elmore 1979; Lipsky 1980/2010; Sabatier 1986; Tummers and Bekkers 2014). This perspective views it as a crucial advantage that "multi-level governance allows decision makers to adjust the scale of governance to reflect heterogeneity" (Hooghe and Marks 2003: 236). Accordingly and as the above quote by Elmore also illustrates, "if local implementers are not given the freedom to adapt the program to local conditions it is likely to fail" (Matland 1995: 148). A bottom-up view thinks of member states who customize EU policy as especially eager problem-solvers. Bugdahn (2005: 182) refers to this phenomenon as "domestication" leading to a progressive interpretation of EU law. She finds that "some forms of domestication (…) clearly had a positive impact on the availability of environmental information" (ibid: 193). From a bottom-up view, we would hence associate extensively customized policies with more successful policy implementation.

Apart from the discussion surrounding gold-plating, the Europeanization literature has hitherto

fallen short of explicitly theorizing, let alone analyzing the empirical consequences of discretionary transposition for the practical application and outcomes of EU policies. Still, existing studies do provide clues on factors affecting the successful practical application of EU law (see Treib 2014 and Knill 2015 for recent overviews). In what follows I combine these insights with Sabatier and Mazmanian's (1980) seminal framework of the implementation process, which continues to inspire contemporary implementation studies (recently, Exadaktylos and Zahariadis 2014; Hinterleitner et al. 2016).

Conditions for successful implementation

Sabatier and Mazmanian (1980) understand the implementation of public policies to be the result of the interplay between three main sets of factors: problem tractability, the ability of a policy to structure implementation, and a favourable political and socio-economic context. One major weakness of Sabatier and Mazmanian's framework is that they did not account for "adaptive" implementation, that is, the use of discretion (Sabatier 1986; see also Matland 1995). To correct for this, I integrate customization levels as a first important condition, and expect it to be relevant, in interplay with other factors, for successful implementation.[1]

Tractable problem (TRACT)

The tractability of the problem at hand refers to the fact that some social problems are easier to deal with than others. This entails questions such as whether a clear understanding exists about the required behavioral change, the size and variety of the target groups, and the extent of

---

[1] I do not integrate misfit arguments even if they may be most plausible in relation to practical application (Treib 2014) because the customization condition already captures the relative distance between EU and domestic policies. Dörrenbächer and Mastenbroek (2017) show that preferences for preserving the status quo matter for the degree to which transposing actors grant discretion to practical policy implementers. Furthermore, the dependent variable of goodness of fit arguments is domestic policy change (Knill 1998), while this paper seeks to explain successful implementation instead. Whether or not successful implementation may or may not require domestic change is not subject of my analysis.

behavioral change required. The greater the extent of behavioral change required and the bigger and more diverse the target group(s), the more challenging a program's successful implementation becomes (Sabatier and Mazmanian 1980: 544). A higher number of target groups can induce distributional conflicts and divergent preferences of involved actors, introducing conflicting interests, veto points and the necessity for compromises affecting different stages of the implementation process (Dimitrova and Steunenberg 2013; Knill 2015; Treib 2014).[2] The condition "tractable problem" captures both the extent of required behavioral change and the complexity of the set of addressees targeted by the rule (number of target groups). I distinguish minor, more tractable micro-issues from more major, macro-issues to account for the extent of required behavioral change.

The ability to structure implementation refers to the questions whether the policy indicates the problem to be addressed, stipulates the objectives to be pursued, selects adequate implementing institutions, provides them with legal and financial resources, influences the policy orientations of implementing agents, and integrates societal actors into the implementation process. Three conditions capture this aspect: a centralized implementation structure, an active enforcement system, and an externally and internally coherent policy design.[3]

Centralized implementation structure (CENT)

The practical implementation of EU law takes place at levels of governance lower than the nation state (Gollata and Newig 2017; Hooghe et al. 2010). This implies that the institutional design of implementation matters (Knill 2015). It is a central insight of multilevel governance

---

[2] In addition, the relative importance or salience of an issue is an important influence on implementation processes (Spendzharova and Versluis 2013; Versluis 2003, 2007). However, I do not account for issue salience in this analysis of practical implementation. First, the allocation of attention to policy issues has partly already explained the extent to which EU rules were customized (Thomann 2015). Second, at the level of practical application, issue salience mainly influences the degree to which EU rules are enforced (Versluis 2003), a factor which I account for separately.

[3] I do not account for administrative capacity in my assessment. While this factor has proven important driver of legal compliance, "the actions of implementers are more influenced by the effectiveness of domestic enforcement and judicial systems than general levels of administrative capacity" (Zhelyazkova et al. 2016: 15).

research that how power and competencies are allocated away from central authority influences the efficiency and performance of policy implementation (Biela et al. 2013; Hooghe and Marks 2003; Skjaerseth and Wettestad 2008). Inherent trade-offs exist between local flexibility and national control, especially when it comes to enforcement (Whitford 2007). Next to the "right to decide" of regional governance units (Federalism), what matters for successful implementation is the regional "right to act", that is, decentralization (which can be pronounced also in non-federal countries; see Keman 2000). I understand a decentralized implementation structure here broadly as one in which the competence to independently implement policies as disposed by some superordinate institution are redistributed or dispersed away from a central location or authority (Biela et al. 2013; Hooghe and Marks 2003; Whitford 2007). Reflecting the top-down versus bottom-up debate discussed earlier, the literature features divergent arguments about the role of decentralization for successful implementation (Biela et al. 2013; Keman 2000; Knill 2015: 23).

To answer the research question of this study, two aspects of decentralization need to be accounted for. The first refers to regional self-rule, that is, the general extent, scope and character of authority of regional governments in their own terrain. This is operationalized "as the extent to which a regional government has the authority to act autonomously, the scope of its policy competencies, its capacity to tax, and the extent to which it has an independent legislature and executive" (Hooghe et al. 2010: 14). The second, sector-specific aspect of decentralization relevant here entails the question of how integrated the implementation structure is (Knill 2015: 22). The existence of several different administrative units implementing and enforcing the specific policy at several levels may enhance the need for coordination and the number of possible "veto points" for a successful implementation (Sabatier and Mazmanian 1980: 546).

Active enforcement system (ENF)

A crucial condition making successful implementation more likely is that the process is legally structured to enhance compliance by both implementing officials and target groups (Sabatier 1986). Monitoring and enforcement have a crucial mediating function especially for the practical application of EU law (Bugdahn 2005; Gulbrandsen 2011; Jensen 2007; Knill 2015; Tosun 2012; Versluis 2003, 2007). Treib (2014: 29-30) highlights two different ideal-typical logics of control systems, drawing on McCubbins and Schwartz' (1984) distinction of congressional oversight of executive agencies. In my analysis, "police patrol oversight" translates into an *active enforcement system* by way of public inspections. This monitoring strategy tends to maximize the temporal and spatial density of controls on farms and veterinary dispensaries in order to detect implementation deficits in the actual procedures of dispensing and administering veterinary drugs, before they become consequential. Contrary to this, more *passive enforcement approaches* us a "fire alarm" logic, that is, an effective and rigorous system through which violations are reported and sanctioned once they occur. In the food safety area, passive strategies emphasize punctual, unannounced controls of reasonable suspects, and/or the sampling of food products in slaughterhouses and supermarkets (rather than in sites of primary livestock production), with a focus on detecting violations of maximum residue limits and pathogens. For a regulatory policy, an active control system is expected to be more effective (Treib 2014).

Coherent policy design (COH)

Many implementation problems are the result of deficient policy design (Knill 2015: 15). Policy design refers to the calibration of policy goals and means at different levels of abstraction (Howlett 2009; Howlett and Rayner 2007). Successful implementation is more likely when objectives are unambiguous, clear and consistent, and the causal theory about how policies and instruments achieve the behavioral change is sound (Knill 2015: 15; Sabatier and Mazmanian 1980: 546; see e.g. Dörrenbächer 2017). In this vein, the condition "coherent policy design"

draws on concepts from the policy evaluation literature (Bussmann et al. 1997; Knöpfel et al. 2011). It consists of a combined assessment of the external and internal coherence of the national policy design (detailed assessments in Sager et al. 2011). External coherence refers to the question whether the veterinary drugs regulations are logically consistent and compatible with the national regulations in neighboring, relevant regulatory fields. Internal coherence assessed in terms of the a) clarity, b) completeness and c) coherence of the national rules with overarching elements of the program.

Table 1: Explanatory framework and directional expectations

Condition	Ceteris paribus, condition produces successful implementation (SUC) when…	Ceteris paribus, condition produces unsuccessful implementation (suc) when…
Adaptive implementation Extensive customization CUST	No expectation	No expectation
Problem tractability Tractable problem TRACT	High	Low
Ability to structure implementation Centralized structure CENT	No expectation	No expectation
Active enforcement system ENF	Present	Absent
Coherent policy design COH	Coherent	Incoherent
Political support Domestic resistance RES	Low	High

Note: directional expectations denote counterfactual arguments rather than empirically testable hypotheses (Schneider and Wagemann 2012: 168-177).

Domestic resistance (RES)

Implementation is not a merely technical, but also a fundamentally political process. Hence, next to the legal structure and the nature of the problem, practical implementation dynamics are also driven by the preferences and strategies of the involved actors (Knill 2015: 26; e.g.,

Beugelsdijk and Effinger 2005; Dörrenbächer 2017; Dimitrova and Steunenberg 2013). A particularly important condition for successful practical implementation is the political support of target groups and policy implementers (Sabatier and Mazmanian 1980: 574; Zhelyazkova et al. 2016). Domestic resistance means that at least one of main target groups with some power to influence domestic policy-making with resources and/or lobbying activities opposes the EU policy.

Table 1 resumes the six conditions for successful policy implementation and outlines expectations about their isolated effects. However, as Knill (2015: 28) highlights, "under what conditions the implementation of policies can be characterized as rather effective or ineffective, can hardly be answered by a mono-causal explanation, but needs to take into account a complex configuration of different aspects".

Conceptualizing and measuring successful implementation

As Marsh and McConnell (2010) outline, what constitutes policy success is itself a highly contentious issue – the answer lies in the eye of the beholder (Marsh and McConnnel 2010; Matland 1995). In this paper I focus on programmatic success, rather than process success (legitimacy of decision-making) or political success (popularity and political usefulness of policy). This focus implies an interest in the question whether EU policies succeed in resolving the policy problems that they are designed to address in practice (Scharpf 1997). Successful implementation in a programmatic sense is about goal achievement (e.g., Skjaerseth and Wettestad 2008). Multi-level systems like the EU additionally introduce a spatial dimension to this question (Marsh and McConnell 2010: 577). As different political systems of cultures imply different ideas of successful implementation, "'one-size-fits-all solutions are often neither

politically feasible nor normatively desirable" (Falkner et al. 2005: 1).[4]

As Thomann and Sager (2017) discuss, the question of what constitutes successful implementation is approached very differently by top-down and bottom-up implementation perspectives (Matland 1995). On the one hand, 'conformance implementation' refers to the degree to which the centrally decided blueprint is implemented from top to down (Barrett and Fudge 1981). This top-down school, which dominates Europeanization research, is primarily interested in comparing the intended and actually achieved outcomes of implementation, where the degree of the goal attainment serves as an indicator for implementation success (Knill 2015).Implicitly or explicitly, top-down perspectives tend to view discretion and the resulting deviations from the centrally decided rule as a control problem (Thomann *et al.* 2016). Alternatively, 'performance implementation' denotes whether a policy achieves outcomes that resolve the original policy problem at stake (Barrett and Fudge 1981). Ultimately, effective implementation is measured by the extent to which the perceived outcomes correspond with the preferences of the actors involved in the implementation process (Knill 2015). From this perspective, diverse approaches of problem-solving are actually an intended result of the decentralized implementation structures of multi-level systems.

To conceptualize these differing ideas into a comprehensive measure of successful implementation, a closer look at the implementation process is warranted. The Europeanization literature commonly divides the implementation process into transposition, practical application and enforcement (e.g., Falkner et al. 2005; Zhelyazkova et al. 2016). However, the distinction between and the sequence of these phases are not as clear-cut as it may appear. For

---

[4] It can be assumed that the procedural, programmatic and political dimensions of policy success interact with each other. As Héritier (2016: 17) points out, 'empirical findings (…) indicate that in public opinion good governance of the state is a more important source of democratic legitimation than the correct and fair democratic procedures as such'. If this is so, customization dynamics, if causally relevant for programmatic success, could be an important source of legitimacy for the EU (Thomann and Zhelyazkova 2017).

example, Knill (2015) distinguishes formal transposition from practical transposition. The former depicts legal and administrative provisions for the transposition of EU law into the national legal and administrative system. Practical transposition, in turn, refers to national regulation practice, that is, the practical application of and adherence to guidelines. Bondarouk and Liefferink (2016, see also Bondarouk and Mastenbroek 2017) include these latter aspects of practical transposition into their measure of practical implementation performance, as well as monitoring activities which are often attributed to the enforcement phase. Versluis (2003, 2007) distinguishes EU law "on paper" and "in action" and includes regulatory inspection activities – as a core aspect of enforcement – into the latter. Tosun (2012) suggests to use organizational inputs (that is, the legally defined competences of agencies to monitor and enforce laws) to measure enforcement activities – which would arguably count as legal transposition according to Knill (2015).

In order to introduce more clarity, I adopt an evaluation perspective and structure implementation processes in view of the involved actors' role in relation to the underlying policy problem. Doing so is not only useful to integrate, to a degree, both top-down and bottom-up conceptions of successful implementation, but also to tackle the "eye of the beholder" problem and get to a clearer idea of different aspects of successful implementation (Marsh and McConnell 2010).

The policy evaluation literature commonly identifies three types of results (that could or could not be achieved) of implementation processes (Knill and Tosun 2012; Knoepfel et al. 2011). *Policy outputs* cover all decisions (on paper) and activities (in action) involving the implementation and enforcement of policy measures by implementing agents (usually administrative agents). In order to resolve the technical problem that underlies a policy, policy measures seek to make certain societal actors change their behaviour – the addressees (or target group, or regulatees) of the policy. The changes in the behaviours of target groups in response

13

to the policy outputs are the *policy outcomes*. One policy output can have several target groups with different corresponding policy outcomes; several outputs often target the same outcome. "Intervention hypotheses" are assumptions about how policy outputs will lead to policy outcomes (Knoepfel et al 2011). The beneficiaries of a policy are then those actors who are affected by the problem and hence benefit (or suffer) from this (lack of) behavioural change. Accordingly, *policy impacts* refer to the effects triggered on the beneficiaries. Assumptions about how policy outcomes achieve policy impacts are called "causality hypotheses" (Knoepfel et al. 2011). Figure 1 illustrates the involved actors, outputs, outcomes and impacts, using the earlier example of the rules for antibiotics.[5]

Figure 1: Policy outputs, outcomes and impacts: an example
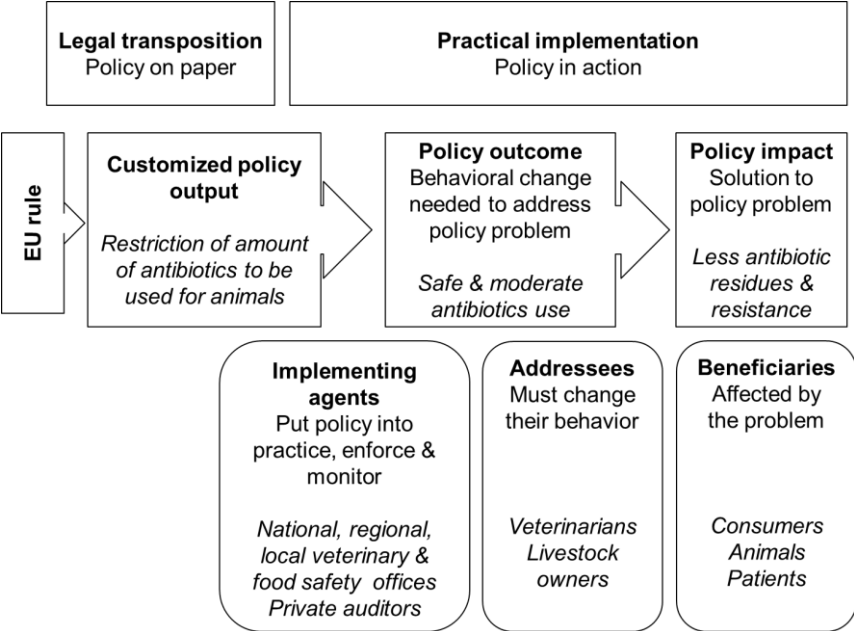
Source: own illustration.
Italics: example (non-exhaustive).

---

[5] It is worth noting that some authors reverse the terms "outcome" and "impact" (e.g., Knoepfel et al. 2011). While I otherwise use the encompassing evaluation framework by Knöpfel et al. (2011), I follow Sager and Rüefli (2005) in using the internationally established terminology (such as Patton 1997, Sabatier and Mazmanian 1980).

This conceptualization allows for several clarifications. First, an evaluation of the effects of customization (or gold-plating) can differ depending on the actor concerned by it. For example, the discussion surrounding the discretionary implementation of EU policies under the heading of "gold-plating" often confuses targets with beneficiaries. The EU Commission's High Level Expert Group on Monitoring Simplification for Beneficiaries of ESI Funds primarily refers to red tape for businesses when it defines gold-plating as "the extra requirements and administrative burden imposed on Beneficiaries by national and sub-national authorities beyond those deriving from provisions at EU or national level". [6] However, higher burdens on businesses as target groups can actually lead to better impacts for beneficiaries; for example, in the field of reducing environmental pollution. We hence need to distinguish outputs, outcomes and impacts in order to understand those effects (e.g., Knill et al. 2012).

Second, we can think of decentralized implementation by member states as a first outcome of an EU policy from the perspective of the EU. Sager and Rüefli (2005) show that multi-level implementation systems add another layer of complexity to implementation processes: the decentralized implementing agents actually become the first addressees of the policy for the centralized policy maker. In the following I will adopt this perspective because it is implicit in most of the EU implementation literature. Hence, I count domestic outputs in action ("practical transposition" with Knill 2015) as a part of successful implementation. Finally, I follow Knill's (2015) suggestion to use policy outcomes, but not impacts, to evaluate the successful implementation of EU policies. As the focus shifts toward policy impacts, it becomes increasingly difficult to establish a causal link with a policy. Numerous factors can affect policy impacts, and (comparable) data availability often poses serious challenges (Knill 2015).

---

[6]

https://ec.europa.eu/futurium/en/system/files/ged/hlg_16_0008_00_conclusions_and_recomendations_on_goldplating_final.pdf (retrieved 17.4.2017).

Even for domestic policy outcomes, empirical evidence is not always easy to find (Toshkov 2012; Versluis 2007). Still, the assumption that a link with policy outputs exists – that is, an intervention hypothesis – lies at the heart of the EU implementation research and hence requires empirical assessment (e.g., Beugelsdijk and Effinger 2005; Falkner et al. 2005; Knill et al. 2012; Toshkov and de Haan 2013; Versluis 2007; Zhelyazkova et al. 2016). One factor that makes it very difficult to establish such a link is the fact that comprehensive data about the outcomes of EU policies is very costly to gain (Hartlapp and Falkner 2009; Treib 2014: 29; see Zhelyazkova et al. 2016 for a pioneering example). This said, the EU does engage in evaluative activities in member states, resulting in reasonably comparable, albeit imperfect data about member state implementation and enforcement of specific EU rules (Mastenbroek et al. 2016; Toshkov 2012). In the area of food safety, the European Commission's Food and Veterinary Office (FVO) performs regular audits of the implementation of EU veterinary drugs legislation, including all three Directives under question here, as well as the non-EU member Switzerland.

Generally, the available evidence about the practical implementation of EU policy follows a deficiency-oriented logic. Reports like these do not comprehensively report national implementation patterns, but focus on reported or observed implementation problems instead (Toshkov 2012). For my analysis, this means that successful implementation is operationalized as the absence of evidence for problems, with an inevitable focus on compliance issues. The kind of information provided in the reports does not allow me to move beyond a trichotomous operationalization of the two indicators, with problems either explicitly reported not to exist (0), or explicitly reported to exist (1), or no indications being present that a problem exists (NA).[7]

---

[7] As outlined earlier, I operationalized policy outputs and outcomes with indicators whose (non-)existence was explicitly mentioned in at last one country. This procedure ensures that values of NA do not simply indicate missing data due to incomplete sources (Goertz and Mahoney 2004). Rather, they indicate that the FVO and the sources consulted by Sager et al. (2011) did not find indications for problems regarding this output / outcome in their audits / evaluations, though in principle looking out for them. Reasons for this could be, for instance, that no

Based on these elaborations, it is now possible to precisely conceptualize successful implementation. It should be clear now that customization refers to the use of discretion at the stage of legal transposition, that is to say, national policy outputs *on paper*. I assess the effects of customization on two policy outcomes: first, on national implementation *in practice*, that is, the domestic policy outputs in terms of practical transposition (administrative activities, including monitoring and enforcement). This aspect captures practical conformance as a first policy outcome for the EU. Second, I assess policy outcomes in terms of the behavioural change of the policy addressees ("practical application"). This aspect is more oriented toward performance, meaning a change in the behaviour that causes the policy problem.

Table 2: Successful implementation as a 6-value fuzzy set

		Outcome		
		Problems absent	*No problems found*	*Problems present*
Output	*Problems absent*	A Fully successful implementation (1)	B Mostly but not fully successful (0.9)	C More unsuccessful than successful (0.3)
	No problems found	D Fully successful implementation (1)	E More successful than unsuccessful (0.7)	F Mostly but not fully unsuccessful (0.1)
	Problems present	G More successful than unsuccessful (0.7)	H More unsuccessful than successful (0.3)	I Fully unsuccessful implementation (0)

Own conceptualization based on Ragin (2009: 91).

Degrees of policy success emerge from combining the results of domestic outputs and outcomes, see Table 2. Clearly, implementation is fully unsuccessful if problems are reported

---

problem exists; or that problems with output delivery mean that data on this policy outcome are not collected, which could however reveal a problem if it was done; but the reason for NA is not an incomplete consideration of data sources by the researcher.

both at the levels of output and outcome (quadrant I). Conversely, implementation is fully successful if both practical transposition and policy outcomes are explicitly reported to be achieved (quadrant A). For the grey zones in between, however, we need to consider the sequence of outputs and outcomes for addressing policy goals, and the fact that outcomes can be "decoupled" from outputs (Falkner et al. 2005; Versluis 2007; Zhelyazkova et al. 2016). The conceptualization follows three principles. First, conclusive evidence about problems or their absence is a stronger indicator of (non-)success than the lack of evidence about problems or their absence. Second, it is of less concern if problems exist at the level of outputs, if the desired behavioural change of target groups nevertheless takes place. Hence, evidence about the absence of problems at the level of outcomes can compensate to some degree for problems with outputs or lacking evidence thereon. Conversely, third, it is of little use if implementing agents deliver outputs correctly or no problems were found here, if that demonstrably does not induce the desired behavioural change of target groups. Thus, conclusive evidence about problems at the level of outcomes always indicate that implementation is more unsuccessful than successful.[8]

Table A1 in the appendix lists the domestic outputs and outcomes for all 19 EU rules analysed here. Table A2 in the appendix provides an overview of the kind of problems encountered at the level of outputs and outcomes, and indicates their frequencies. I discuss examples in the results section.

---

[8] Bondarouk and Liefferink (2016) and Bondarouk and Mastenbroek (2017) have recently proposed an alternative, sophisticated conceptualization and measurement of national and subnational differences in implementation performance in practice. I do not adopt this proposal in this analysis for two reasons. The first, theoretical reason is that my analysis adopts a cause-effects logic as used in the policy evaluation literature, rather than a descriptive perspective. This evaluative perspective also requires me to move beyond policy outputs, contrary to Bondarouk and Mastenbroek (2017). The second and intertwined, empirical reason is data availability. The FVO audits that form my database adopt a problem-focused logic rather than providing a comprehensive picture of national and subnational implementation. However, future research should definitively look into the effects of customization using these authors' more comprehensive conceptualization of the substance, scope and effort dimensions of implementation performance.

Data and methods

In what follows I outline the rationale of case selection, the method, as well as the data sources and decisions about measurement and calibration.

Case selection

Veterinary drug regulations are both under-researched and a particularly illustrative example of positive integration. By preventing and managing animal diseases and antibiotic residues in food, the regulations ensure animal health and food safety across borders. Ensuring the exportability of their food products is essential for the countries, who thus generally comply with the EU directives (Sager et al. 2014). Veterinary drugs are dispensed by veterinarians or pharmacies to the end users – veterinarians or livestock farmers – who administer the drug to the livestock. The domestic nature of these processes makes standardization unlikely and customization likely. The domestic regulations of single issues regulated by the EU serve as the units of analysis. This entails 13 EU dispensing rules and 6 EU administration rules, which are a) regulated in an EU directive, b) not instances of full standardization, and c) distinguishable from other processes. The EU policies (10 flexible and 9 inflexible instruments) stem from three directives: Council Directive 90/167/EEC on medicated feedingstuffs, Directive 2001/82/EC on veterinary medicinal products and Commission Directive 2006/130/EC on the prescription requirements. Table A1 in the appendix summarizes these policies and their EU legal basis.

I compare the domestic regulations of Austria, Germany, France, the UK and Switzerland in 2011 (N = 95). A most-similar-cases design was employed to preclude the risk that contextual features explain the observed differences (Blatter and Haverland 2012: 42). These countries share a tendency toward low compliance (Falkner et al. 2005), similar regulatory contexts and functional problems. Because livestock farming has a similar relevance for agriculture (Sager

et al. 2011: 301), the countries face comparable regulatory requirements. Food safety scandals triggered the relevance of veterinary drug regulations, and livestock farming has a similar significance for agriculture: the value added to the gross domestic product by agriculture was between 1 and 4.7, but below EU average in 2000.

Considering a non-member state enables me to assess whether the role of customization for successful implementation differs between EU member states (where EU rules were transposed) and non-member states (where existing national rules are mostly adapted to EU legislation to ensure equivalence). Since 1 January 2009, Switzerland has a contractual obligation for equivalence to relevant EU rules (veterinary agreement). As a consequence, its legislation is checked for Euro-compatibility (Linder 2011: 46). This can lead to the transposition of EU law into Swiss law, however often subject to the regular legislative and democratic processes (Maggetti et al. 2011). The Swiss administration tends only to identify a need for revision if a rule directly contradicts the EU provision. This makes Switzerland a particularly interesting case for assessing how highly national-specific rules relate to successful policy implementation. All domestic regulations analyzed here were subject to extensive revisions since the EU directives were issued (Sager et al. 2011: 301-302; Sager et al. 2014). The EU rules had generally been adopted by all five countries by 2011 (Sager et al. 2011). There are four exceptions: Switzerland did not yet have equivalent rules to the EU in 2010 regarding the documentation of dispensing and administration, as well as the amount of veterinary drugs and medicated feed that could be dispensed.

Although covering a variety of non-EU member countries might enhance the results' analytical leverage, Switzerland is the only European non-EU member that complies with the "most similar" assumptions (see Table A4 online appendix). Functionally, the European non-EU member's similar agricultural sectors should have a comparable significance; a high significance of fishery, for instance, would create entirely different regulatory requirements.

Relative democratic and political stability since the early 1990s when the first EU regulations emerged should ensure the country's comparable regulatory capacity.

Method

The choice of the method follows two core features of my theoretical argument (Toshkov 2016). First, Sabatier and Mazmanian's (1980: 554) model should be understood as a "minimum list of crucial conditions", rather than as individual factors which work in isolation. Second, they present their framework as entailing jointly sufficient and individually necessary conditions for successful policy implementation (cf. Sabatier 1986: 23). Accordingly, I employ Fuzzy Set Qualitative Comparative Analysis (fsQCA) to identify necessary and/or sufficient (combinations of) conditions for an a successful or unsuccessful implementation of EU rules (Ragin 2008; Rihoux and Ragin 2009; Schneider and Wagemann 2012 (software: R packages QCA and SetMethods; Dusa 2007; Medzihorsky et al. 2017). Contrary to other methods, QCA models this kind of causal complexity, which has three elements. *Conjunctural causation* indicates that the theorized conditions affect successful implementation in combination rather than in isolation, a central assumption of Sabatier and Mazmanian's (1980) framework. *Equifinality* means that various configurations of conditions can lead to successful policy implementation, depending on the context (a central finding of EU implementation research, see Falkner et al. 2005). Lastly, *asymmetrical causation* means that the change in the outcome is different when the condition is present as when it is absent. Accordingly, a successful policy implementation can be the result of very different combinations of configurations than unsuccessful implementation.

Given that this method is relatively rarely applied in political science, I shortly explain its rationale (for detailed descriptions, see Rihoux and Ragin 2009; Schneider and Wagemann 2012). The set-theoretic method focuses on configurations of variables as sets in which cases

have membership or not. Set membership requires a statement about a qualitative state: cases are either (more or less) *in* a set, or (more or less) *out* of a set. The attribution of cases to sets is called calibration. Fuzzy sets allow us to account for differing degrees to which phenomena, such as customization or the coherence of the policy design, are present. This approach takes "into account the fact that most social science concepts establish qualitative differences between cases in principle, but that cases manifest adherence to these criteria in various degrees" (Schneider and Wagemann, 2012, 16). Qualitative anchors determine the stage at which the condition is deemed fully present (fuzzy value ≥ 0.95), fully absent (fuzzy value ≤ 0.05) and an indifference (or crossover) point at 0.5. Contrary to usual measurement scales, the crossover point establishes the difference in kind. For example, fuzzy membership values in the set 'successful implementation' above 0.5 means that implementation was rather or fully successful (SUC), while values below 0.5 indicate that implementation was rather or fully unsuccessful. The absence of a phenomenon (i.e., the negation of a set) is indicated by lowercase letters (suc).

FsQCA is based on the fuzzy extension of Boolean algebra. It uses the logical operators OR (+) and AND (*). The logical AND depicts combinations of conditions, subsequently called configurations or paths. A condition is necessary for a certain outcome if the outcome (usually) does not occur in the absence of this condition. The analysis of necessity (super-/subset procedure) starts with simple conditions that are necessary for the outcome (here: equal or unequal performance). If no simple condition proves necessary, further simple conditions can be added using the logical OR until necessity is obtained. I only interpret a condition as necessary that passes the consistency threshold, is not empirically trivial ('unsurprising'), and can be considered meaningful considering conceptual and theoretical knowledge (cf. Schneider and Wagemann 2012; see thresholds in key to Table C1).

When a condition (typically) results in a certain outcome, it is sufficient. For the analysis of

sufficiency, a 'truth table' is constructed. The rows of the truth table indicate all possible combinations of conditions. This enables me to attribute the cases accordingly to the truth table and identify empirically unobserved configurations (so-called logical remainders). If all or enough cases' fuzzy set membership in a truth table row is smaller than or equal to its membership in the outcome, then the row is identified as a sufficient path for the outcome. Subsequently, the logical minimization process identifies the shortest possible expression depicting the combinations of factors that imply ($\rightarrow$) the outcome – the solution term. This is a straightforward procedure that relies on basic set theory: For example, if we observed both TRACT*COH*RES and TRACT*COH*res $\rightarrow$ SUC, then whether or not the policy faces domestic resistance obviously is not relevant. We can reduce this accordingly to the statement that the combination of a tractable problem with a coherent policy design was sufficient for successful policy implementation: TRACT*COH $\rightarrow$ SUC.

In order to evaluate our results, I use consistency and coverage measures. The values of these fit indices can range from 0 (low) to 1 (high) (Schneider and Wagemann 2012, 128). *Consistency* is the extent to which the results are in line with the statements of necessity or sufficiency. This relationship is weakened by 'deviant cases consistency in kind': cases with different membership in kind in the condition set and the outcome set. For sufficient conditions, consistency is indicated for single truth table rows (raw consistency), for single configurations of, or for the whole solution term. Furthermore, the proportional reduction in inconsistency (PRI) indicates the degree to which a given configuration is not simultaneously sufficient for both the occurrence and the non-occurrence of the outcome. Consistency should not be below 0.75 for sufficient conditions, and not below 0.9 for necessary conditions.

For sufficient conditions, *coverage* depicts how well the model explains the available empirical information. Raw coverage expresses how much a single configuration covers, and unique coverage indicates how much it uniquely covers. Low coverage means that the model has a

limited capacity to explain the outcome. For necessary conditions, coverage expresses their relevance in terms of the condition set not being much larger than the outcome set, and the relevance of necessity (RoN) in terms of the condition being close to a constant (all formulae in Schneider and Wagemann 2012,128, 139, 235-239).

I apply the Enhanced Standard Analysis (ESA) procedure and interpret the intermediate solution. This implies, first, that I make theoretically informed, counterfactual assumptions about empirically unobserved truth table rows (logical remainders), see Table 1. Second, I ensure that the coding of the outcome in the truth table does not contradict prior findings of necessity or sufficiency (for a detailed description of this procedure, see Schneider and Wagemann 2012, 198-211). The truth tables, directional expectations, complex and parsimonious solution terms, simplifying assumptions and replication codes are all indicated in the appendix. The raw data will be published online.

Data, measurement and calibration

The original case study evidence for this study was collected for the Swiss Federal Office of Public Health by Sager et al. (2011). Methods comprised an analysis of legal documents, policy documents, secondary literature, telephone interviews and written questionnaires with agents of relevant stakeholder groups and the public administration. Complementary to the data gathered by Sager et al. (2011), I use the reports summarizing the results of the FVO audits and formulating recommendations as a basis to code my outcome "successful implementation" through document analysis (see Table A3, appendix).[9]

Table 3 illustrates the measurement and calibration of the outcome and the six conditions. Calibration decisions importantly influence the results and should be well-reasoned, transparent

---

[9] Retrieved from http://ec.europa.eu/food/audits-analysis/audit_reports/index.cfm (last accessed on 12.5.2017).

Table 3: Measurement and calibration

Set	Operationalization	Calibration anchors				
		0	0.33	0.5	0.67	1
Successful implementation (SUC)	Composed set of 2 variables: domestic policy outputs and policy outcomes, for each implementation problems either explicitly being reported to be absent (0), to be present (1), or no evidence for problems being reported (NA)[1,3]	Six-value fuzzy set, see Table 2				
Extensive customization (CUST)	Added index (0 - 4) of additional *density* and *restrictiveness* of domestic regulation as compared to EU directive, each ranging from 0 (absent) over moderate (1) to extensive (2)[1]	0	--	1.5	--	4
Tractable problem (TRACT)	Composed index (1-6) of micro-issues (0) versus macro-issues (3) and the number of target groups (1-3)[1]	6	5	--	3-4	1-2
Centralized implementation structure (CENT)	Combines 2 indicator sets: CENT = reg + INT *Strong regional self-rule (REG)* Index of regional self-rule in 2006[2]	7.75	--	9.5	--	15
	Integrated implementation structure (INT) Number of institutions involved in monitoring and enforcing the respective rule in a given unit (e.g., a veterinary dispensary or a farmer) at different organizational tiers[1]	11	10	--	9	8
Active enforcement system (ENF)	More police-patrol or more fire-alarm logic of control system, separately for administration and dispensing policies[1]	Fully passive	Legally active, de facto passive	--	Public passive, private active	Fully active
Coherent policy design (COH)	Fully coherent (1) or partly or fully incoherent (0) for six aspects: external coherence; suitability of enforcement structures for service provision; internal coherence of problem definition, objectives, functional, and organizational parameters	2	--	--	--	4
Domestic resistance (RES)	Added index of opposition of target group (yes = 3, no = 0) and its power to exert influence (absent = 1, medium = 2, significant = 3)[1] - Please indicate 2-4 interest groups that are influential in the formulation of veterinary drugs regulations? - How would you rate their power to exert influence (networks with the public administration, political relevance, activities such as, and/or resources for, lobbying at national and European level)?[1]	1-2	3	--	4-5	6

Sources:
[1]Sager et al. 2011. [2]Hooghe et al. 2010. [3]FVO audit reports, see Table A3 appendix.

and based on theoretical and empirical case knowledge (Schneider and Wagemann, 2012, 32-35, 287-291). Due to limited space, all decisions, descriptive statistics, and skewness tests are outlined in detail in the appendix.

Results

I begin with a brief descriptive overview of the situation in the five countries. Figure 2 shows how countries score on average on the conditions and the success of implementation.

Figure 2: Country profiles
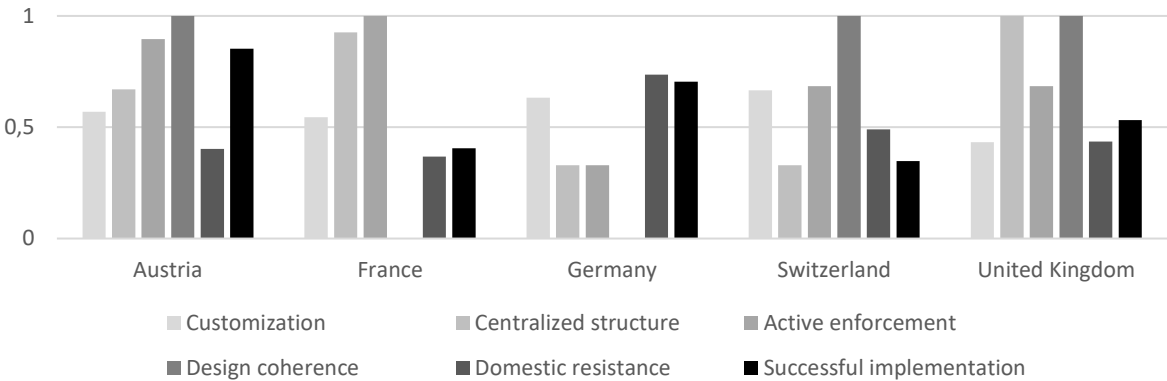
Average set membership of cases sorted by country.   N = 95.
Values above 0.5 indicate a feature's partial to full presence, while values below 0.5 indicate its partial or full absence.

Austria. Austria is a moderate customizer of EU policies, but it has transpsposed some EU rules more restrictively in order to facilitate a controlled dispensing and use of veterinary drugs within a semi-private governance structure of regional animal health services (*Tiergesundheitsdienste*), where relatively far-reaching individual rights for livestock farmers are coupled with ducation requirements and a close collaboration with veterinarians. As a "federation without federalism", Austria's nine regions have no decision-making competences. The governance of the food safety chain is centralized within one independent rgulatory agency.

Its public enforcement strategy is active for dispensing policies, while a fire-alarm strategy on livestock farms is complemented with intensive monitoring activities by vets. Though it does not explicitly define a problem and associated policy goals, the policy design is coherent with neighboring regulations and entails a suitable enforcement structure, functional and organizational rules. Given their close involvement into policymaking, the target groups seem to be content with the regulations. Austria encounters by far the least problems when implementing the EU food safety rules; its output performance is flawless.

France. The French customization strategy closely resembles that of Austria. It has a rather liberal regulatory approach and even some regulatory gaps especially in the area of administration, written agreements between livestock holders and vets being a central instrument for governing the supply and use of veterinary drugs. But France compensates for that with a very active and consistent public enforcement strategy. While French regions have low authority, no less than three ministries share competences in the food safety area, each with their own monitoring and enforcement authorities. Despite the dense enforcement structure and a clear problem definition, the external and intrnal coherence of the policy design are judged as rather low: the operational rules partly clash with the goal of animal welfare, and the regulatory design insufficiently takes into account the scarcity of vetrinarians especially in rural areas, hence raising economic pressures for livestock owners. While the EU rules tend to be rather uncontested, France has a lot of implementation problems: the monitoring and enforcement prove partly ineffective since many controls are announced in advance, are not planned on a risk-based manner, and sanctions are improbable. Furthermore, the regulatory design does little to counter the economic incentives of both vets and producers for an excessive use of antibiotics.

Germany. Germany, in turn, is an extensive customizer. In line with its overarching national strategy to reduce antiobotic resistance, Germany implements considerably more rstrictive rules

for veterinary drugs than minimally required by the EU. In contrast, the regulatory density is rather low especially in the area of administration. The implementation and enforcement of food safety regulation is almost fully the competence of the regions and there is a complex network of local implementing authorities. Due to the size of the sector and often scarce resources, an active enforcement strategy on paper is de facto rather punctually implemented. The German policy design is coherent with other regulations and outlines clear and mutually coherent objectives. However, it lacks a problem definition, the operational rules are extraordinarily complex which affects their clarity. The organization of the sector is so decentralized that enforcement is not uniform. Accordingly, many food safety rules are contested among the addressees. Despite this, comparatively few implementation problems are reported, apart from the sectoral differences in enforcement and a general lack of effective monitoring and federal oversight.

Switzerland. Switzerland turns out to be the most extensive customizer among the five countries, but mainly in terms of regulatory density and differentiation. Particularly, the four Swiss rules that were not yet equivalent to the EU template by 2010 are actually more lenient than the latter. Given that Switzerland only has an obligation for legal equivalence with EU food safety law since 2009, this reflects country-specific, pre-existing regulations. The Swiss regions have high decision-making and implementation authority. Regulatory competences are split between two separate ministeries and partly delegated to private, for profit veterinarians. Switzerland combines an active enforcement strategy in the dispening area with a fire-alarm public control approach at the level of administration. Here, private vets are supposed to jump in, but but they often don't in practice (Sager et al. 2014). The Swiss veterinary drugs regulations are partly inconsistent with the regulations of alternative medicines, and problem definitions are focused mainly on human rather than veterinary medicine. Otherwise, the enforcement structure, goals, operational and organization measures are clear and consistent.

Levels of domestic resistance are intermediate. Yet Switzerland faces many implementation problems. Federal oversight and coordination are low, enforcement activities differ starkly between regions and are often deficient (Thomann 2015b). These problems also translate often into deficient policy outcomes: particularly because the private vets often prioritize their own economic interests over their role as co-producers of food safety regulations (Thomann and Sager 2017c).

United Kingdom. Consistent with its liberal tradition, the UK tends not to add restrictiveness to EU rules. If it customizes them, it is in order to differentiate them or to create exemptions. As avoiding unnecessary regulation is an explicit goal of the overarching "Animal Health and Welfare Strategy", many aspects of food safety are targeted through voluntary codes of good practice. The British food safety regulations generally strongly rely on private regulation to foster these codes, e.g. via so-called farmer assurance schemes. Though the four regions have separate implementing organizations, overall the implementation structure is highly centralized. The use of veterinary drugs is only weakly controlled on farms; instead, using a fire-alarm approach, end products in supermarkets are checked for residues, and violations are then prosecuted. Next to a coherent enforcement structure, the British food safety policies are continuously improved in exchange with stakeholders. Thus, they are externally coherent, formulate clear problems and goals. Still, the functional parametres may be too loose to prevent clashes with the goal of animal welfare, and record keeping requirements are hardly enforced. Especially some more intrusive EU rules face opposition in the UK. However, this system tends to be effective. Implementation problems relate to the output of enforcement. Specifically, there are low levels of coordination between regional authorities, and between local and central authorities, as regards control activities and the exchange of information.

In a first explanatory step, the *analysis of necessity* reveals that there are no necessary conditions for either successful nor unsuccessful implementation (see Table C1 appendix). In a second

step, I look at the *relation between customization and successful implementation*, see Figure 3.

Figure 3: Sufficiency of customization for successful and unsuccessful implementation
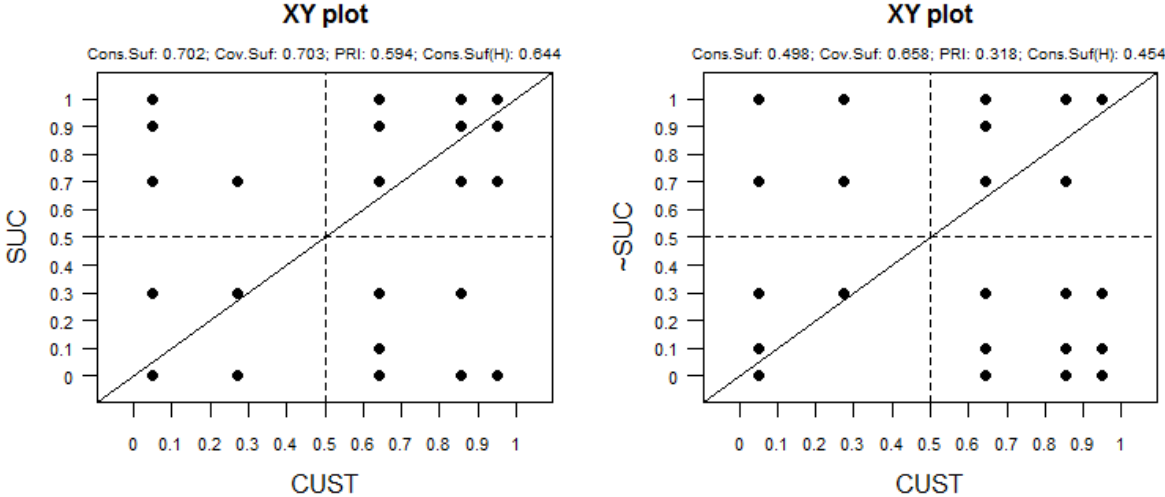
As the Figure 3 shows, the fact that member states customize EU policies extensively is neither sufficient for successful nor for unsuccessful policy implementation on its own. There is also no statistically significant correlation between customization and successful implementation (Pearson's $R = 0.128$, $P = 0.217$). In other words, my cases show no straighforward link between the use of discretion during implementation, on the on hand, and the success of implementation, on the other. Thus, in a third step, I analyze how complex configurations of conditions affect the success of implementation.

Table 4 presents the seven paths that imply a successful implementation of the EU food safety rules (SUC), and another six paths that imply their unsuccessful implementation (suc). Each column represents a sufficient path. Black dots indicate that a condition is present, white dots indicate that a condition is absent, and blank spaces indicate that a condition is irrelevant. The single cases that are explained by this solution, the consistency and coverage indicators for the

Table 4: Sufficient conditions for successful and unsuccessful implementation (intermediate solutions)

	Outcome: successful implementation (SUC)							Outcome: unsuccessful implementation (suc)					
Path	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Extensive customization CUST	●	●	•	○	○			○	○	○	○		●
Tractable problem TRACT		○	●	●	○	●	○		○	○	●	●	●
Centralized structure CENT	○	●	○	○		●	●	●		○	●	●	
Active enforcement system ENF		•	●	○	•	●	○	●	●				○
Coherent policy design COH	○					●	•	○				○	●
Domestic resistance RES	●	○	●	○	●	○	○		○	○	●	●	○
Consistency	0.921	0.888	0.884	0.919	0.883	0.900	0.893	0.801	0.763	0.652	0.764	0.798	0.815
Raw coverage	0.153	0.238	0.198	0.089	0.122	0.266	0.060	0.251	0.246	0.130	0.368	0.206	0.060
Unique coverage	0.061	0.094	0.026	0.021	0.014	0.106	0.025	0.055	0.028	0.015	0.095	0.059	0.020
Cases	d10ge, d2ge,d4ge, d5ge,d7ge; a3ge,a4ge, d3ge,d9ge	d6fr,d7fr; d10au,d10uk ,d1au,d4au, d5au,d7uk	d8ch, d9ch	d11ge; a2ch	a6au	a5au,d11au, d11uk,d3uk, d8au; a1au,a4au, d12au,d13au, d13uk,d3au, d8uk, d9au, **d9uk**	a6uk	d4fr,d5fr; a6fr; a2fr, **d11fr**, d8fr	d4fr, d5fr; d1uk, d4uk, d5uk	a6ch	a2au	a4fr, a5fr	a1ch; a3uk
Solution consistency							0.894						0.773
Solution PRI							0.831						0.584
Solution coverage							0.626						0.625

Black circles indicate the presence of a condition, and white circles its absence. Blank spaces indicate the irrelevance of a condition. Large circles indicate the parsimonious solution (suc: enhanced parsimonious solution). Cases separated with semicolons belong to different truth table rows. **Bold** are deviant case consistency in kind. Raw consistency thresholds and untenable assumptions in appendix C, directional expectations in Table 1.

single paths and the overall solution are listed below. Cases can display several paths. To facilitate the interpretation of the complex configurations, I discuss typical cases for each path.

Whether or not EU policies are customized is relevant in five of these seven constellations. In a majority of the explained cases of successful implementation, these policies regulate the area of dispensing, and they were extensively customized when transposed into national law.

In **path 1**, covering nine dispensing policies in Germany, the EU rules faced domestic resistance, the implementation structure is decentralized, and the policy design has incoherent elements. A good example is the policy for the amount of prescription drugs that may be dispensed. As mentioned earlier, Germany defined "the amount needed for one treatment" as being limited to 31 days and even to 7 days for antibiotics. The vague EU wording is contested in Germany as medical, consumer, and organic agriculture associations strongly call for measures to reduce antibiotic resistance. The Germany regulations do justice to these widespread concerns. Though the decentralized enforcement structure produces significant problems in enforcement (outputs), the FVO auditors emphasize that the available evidence suggests satisfactory compliance with these rules by veterinarians and livestock owners.

In **path 2**, eight extensively customized EU dispensing policies in France, Austria and the UK combine with low problem tractability, a centralized implementation structure, an active enforcement approach, and low domestic resistance. A typical case are the dispensing rights of veterinarians and pharmacies exceeding the mere dispensing of drugs in France. According to the EU, dispensing actors need a national authorization if they want to manufacture, pack, bottle, or customize veterinary drugs that are produced on an industrial or commercial basis. France concretizes this rule by allowing for the manufacture of veterinary drugs in a pharmacy both upon prescription by a veterinarian, and according to a dispensatory. The former includes veterinary dispensatories and vaccines that are manufactured on the instructions of a vet from pathogens or antigens obtained from an animal and used for the treatment of that animal or

32

other animals on the same site. These rules are especially relevant for veterinarians since such preparations can be important to ensure an adequat treatment, and constitute an additional source of income. Accordingly, the customized French solution, granting vets controlled options in situations where it is warranted, is uncontested. Given that these rules are uniformly and well-enforced, no problems are reported at the levels of outputs or outcomes.

Path 3 also entails extensively customized dispensing policies, but it covers only two policies in Switzerland about the documentation of dispensing. These are merely administrative measures of who needs to keep the documents for how long; A tractable policy problem meets a decentralized implementation structure, an active enforcement system, and EU rules that are resisted domestically. The reason for resistance is that the EU requires dispensing actors to keep track for at least 5 years, whereas Switzerland has traditionally obligated both dispensing and administrating actors to keep track for only three years. The target groups perceive stricter rules as red tape, though it would be little effort for them (Sager et al. 2014). Despite the regional differences in enforcement, target group compliance with these customized record-keeping requirements is reportedly good.

In a minority of three policies, the fact that EU policies were NOT customized is relevant for successful implementation. All three policies concern relatively inflexible EU standards. In the cases of the cascade rule in Switzerland and prescription forms for medicated feed in Germany (**path 4**), the problem was tractable and the rule uncontested, while the implementation structures are decentralized and enforcement is rather passive. The regulations of withdrawal periods in Austria (**path 5**) – one of the trickiest aspects of ensuring food safety in livestock production – were actively enforced under a coherent overarching policy design. Vets are keen to comply, they complain about the rule because it is unclear to them how withdrawal periods should be harmonized for older products, and defined for bees.

Finally, customization proves irrelevant for successful implementation in two scenarios.

Empirically most relevant is the ideal scenario, covering 14 in Austria and the UK, when the problem is tractable, implementing structures are centralized, the enforcement system is active, the policy design coherent, and domestic resistance is low (**path 6**). Similarly, in the case of withdrawal periods in the UK, compliance is good even in the presence of aless tractable problem and a fire-alarm control system (**path 7**). Experts grant crucial relevance to this system of controlling for residues in end products for this outcome.

Figure 4: Sufficient conditions for successful implementation
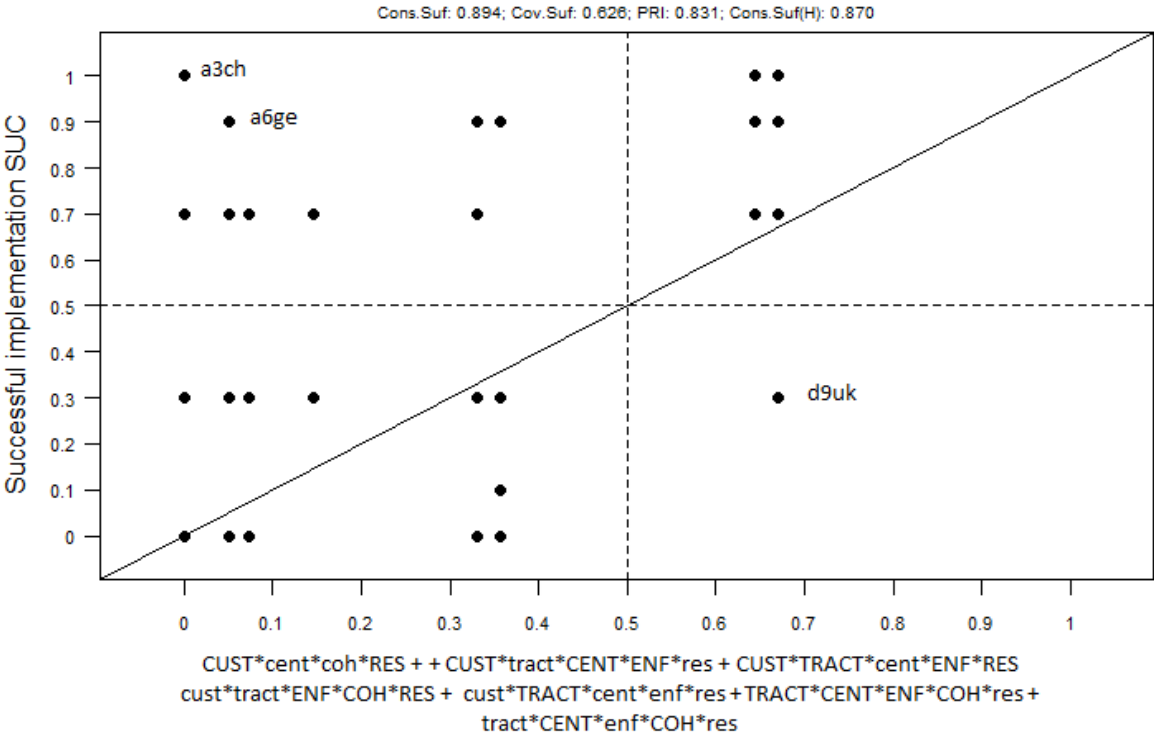
Cases situated above the diagonal are consistent. In the upper left quadrant are deviant cases for coverage, in the lower right quadrant are deviant cases consistency in kind. The lower left quadrant is irrelevant (Schneider and Rohlfing 2013).

Figure 4 shows that this solution has a good consistency, there is only one case (d9uk) that contradicts the pattern. Coverage is acceptable with the solution explaining successful implementation in 36 out of cases (61 per cent). I discuss deviant cases below.

I now turn to the six scenarios that typically result in unsuccessful implementation. It immediately catches the eye that the fact that customization is limited matters in all but four of these cases, and that these are often policies of administration.

Path 8, covering 6 French policies, entails a literal transposition together with a centralized enforcement structure, an active enforcement system, and an incoherent policy design. An interesting example is the cascade rule in France. The goal conflicts with animal welfare are particlarly pronounced in cases of supply shortage: veterinarians either stick to the cascade rule which forces them to obtain a time-consuming import authorization or resort to often unsuitable human medicines, or they also have significant incentives to not comply and simply import suitable medicines illegaly from abroad. As outlined earlier, the way in which the numerous inspections are put into practice renders them ineffective. Together, this leads to significant an uneffective regulation in practice. Accordingly, the stakeholder interviews reveal a *widespread demand for a more customized regulatory solution*.[10]

In **path 9**, the unsuccessful implementation of five dispensing policies in France and the UK can be explained by a combination of lacking customization and problem tractability, an active enforcement system, and low domestic resistance. An interesting illustration is the way in which the EU requirement for a clinical examination prior to prescription (a vet must have made sure the treatment has a medical justification) is transposed very loosely in the UK. De facto, a vet can simply look at the clinical files without an actual examination of the animals if he or she is sufficiently familiar with them to pose a diagnosis. Obviously, the British farmers and vets appreciate that this lack of restrictions makes their life easy in a significant area of their

---

[10] « En conclusion, les acteurs estiment que les mécanismes palliatifs tels que la cascade ou les ATU posent d'innombrables problèmes dans la pratique. Ils attendent la mise en place de solutions de fond au problème de la disponibilité, qui leur permette d'évoluer dans un cadre clair qui ne soit pas marqué par des pratiques au coup par coup. » (Sager et al. 2011 137).

everyday practice. However, due to a lack of coordination between regional and national authorities, these rules are not sufficiently enforced and it is unclear whether they are eventually effective in ensuring a responsible use of veterinary drugs.

Literal transposition also combines with other factors to coincide with unsuccessful implementation when it comes to withdrawal periods in Switzerland (**path 10**), and the cascade rule in Austria (**path 11**). These cases contrast in that problem tractability is low, the implementation structure is decentralized, and domestic resistance is absent in Switzerland, where enforcement problems prevail, whereas the exact opposite holds for Austria, where the EU rule is in effective in improving the availability of treatment options. Here, too, *stakeholders explicitly demand a more customized solution that enables effective and swift problem-solving in supply shortage situations, and provides more detailed information.*[11]

In **path 12**, customization is irrelevant. These two administration policies in France related to tractable problems, were implemented within a centralized structure, while the policy design was incoherent and the rule opposed domestically. Both rules refer to the possibilities of livestock farmers to blend drugs with feed, which is legally allowed but de facto tied with impossibly high regulatory requirements in France. For these policies, the checks that should be carried out by private veterinarians have crucial function in ensuring farmers' compliance. Simultaneously, however, the hybrid policy design gives the vets economic incentives to be lax in this regard. Meanwhile, the public enforcement system also suffers from problems (announced inspections, low probability of sanctions).

Lastly, there are two administration policies for which extensive customization contributed to unsuccessful implementation, combined with a tractable problem, a fire-alarm control

---

[11] „ Tierärzte stellen bisweilen die Stringenz der Umwidmungsregeln in Frage, wenn sie die Lösung zu einem medizinischen Problem kennen, sie aber nicht durchführen dürfen.(…). Tierärzte fordern generell bessere Informationen dazu, was bei einer Umwidmung erlaubt ist und was nicht." (Sager et al. 2011: 237).

approach, a coherent design and low domestic resistance (**path 13**). For example, the EU allows both vets and livestock holders to administer the drugs used off-label. The Swiss regulation restricts this right for livestock holders to two situations: either there is a written agreement with a vet, or a vet supervises the on-farm manufacture of medicated feed. While coherent on paper and popular with the addressees, in practice this hybrid design leads the vets to neglect their control function vis-à-vis their customers, while public oversight also suffers from the above mentioned enforcement deficit. As a result, outputs suffer and it is unclear whether the regulations are actually complied with.

Figure 5: Sufficient conditions for unsuccessful implementation
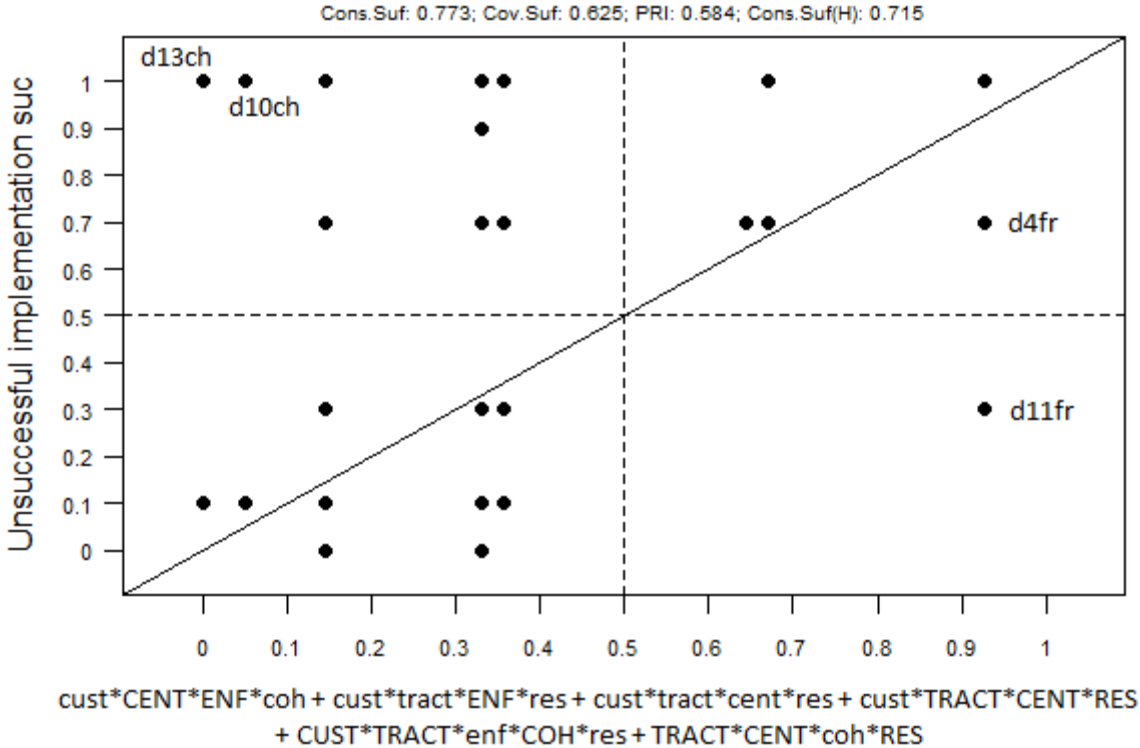
Cases situated above the diagonal are consistent. In the upper left quadrant are deviant cases for coverage, in the lower right quadrant are deviant cases consistency in kind. The lower left quadrant is irrelevant (Schneider and Rohlfing 2013).

Figure 5 illustrates that this solution, too, has a good consistency, with only one deviant case

consistency in kind (d11fr). However, the decent coverage value conceals that the solution explains only 14 out of 36 cases of unsuccessful implementation (39%).

In a next version, the paper will discuss deviant cases for consistency and coverage, based on Table 5.

TO BE WRITTEN

Table 5: Post-QCA case comparisons

Case	Type of case	Comparison with	Question
a3ch	Deviant case for coverage with successful implementation	Case with identical configuration of conditions and unsuccessful implementation: **a4ch**	Which additional condition distinguishes a3ch from a4ch and which fosters successful implementation in a3ch?
d9uk	Deviant case consistency in kind: configuration of conditions should imply successful implementation , but did not	Case with successful implementation, member of similar truth table row: **d11uk**	Which additional condition(s) does d11uk not display and which foster(s) unsuccessful implementation in d9uk?
d13ch	Deviant case for coverage with unsuccessful implementation	Case with the identical configuration of conditions and economic performance: **d3ch**	Which additional condition distinguishes Spain from Portugal and which fosters unequal economic performance in Spain?
d11fr	Deviant case consistency in kind: configuration of conditions should imply unsuccessful implementation, but did not	Case with unsuccessful implementation, member of the same truth table row: **d8fr**	Which additional condition(s) does d8fr not display and which foster(s) successful implementation in d11fr?

Based on Schneider and Rohlfing 2013.

Conclusions

As Knill (2015: 16) highlights, cross-country differences of EU rules that were originally intended to be uniform constitute a form of "renationalization" of EU policies. Scholars of Europeanization have long pointed to the importance of the diversity in policy implementation for the idea underlying the European experience (Héritier 1998; Majone 1999). Simultaneously,

compliance research reminds us of the problematic consequences of a lacking commitment of member states to EU policies. This discourse relates to an old controversy in the theory and practice of policy implementation about the role of discretion for successful policy implementation (Matland 1995). In times in which the legitimacy of European integration is increasingly challenged, it is important to gain a better understanding of the policy consequences of uneven implementation.

It has been the basic premise of this paper that what matters in this discourse is the degree to which the EU is ultimately able to find joint solutions to shared policy problems (Scharpf 1997). In this vein, the paper joins a recent shift in Europeanization research toward a more performance-oriented perspective (Thomann and Sager 2017 a, b). Such a perspective understands implementation performance fundamentally as the result of a process of interpretation of EU law which emerges from an interplay between Europeanization and domestication forces (Bugdahn 2005). Not only do we know very little about the diversity of policy solutions beyond compliance in the EU (see Thomann 2015a); we also know very little about the practical implementation of EU policy beyond transposition (see Zhelyazkova et al. 2016). This paper addresses both gaps. Specifically, this is the first systematic comparative analysis of the consequences of the customization of EU policies during transposition – other than questions of non-compliance – for the practical success of policy implementation in the EU.

Overall, the findings suggest that an extensive customization of EU policies can play an important role for successful implementation in certain situations, and particularly because of how national legislators accommodate the needs of the actors concerned by the rules. Conversely, customization is less relevant for successful implementation under otherwise ideal domestic circumstances. Furthermore, the result suggest that in certain situations, literal transposition often is not beneficial for achieving policy outcomes in practice. In particular,

target groups repeatedly expressed their demand for more customized solutions. Furthermore, some of the evidence suggests a clear positive link of customization with practical compliance; however, this effect dos not occur in isolation, but in interplay with other relevant aspects of the policy, the implementation arrangement, the policy design, and attitudes of the target groups.

This evidence encourages us to rethink the paradigm of avoiding "over-implementation" in the EU regulatory state. Scholars and practitioners in the EU tend to think of deviations going beyond the minimum required by EU law (also called gold-plating) as a problematic outcome: it leads to red tape and creates unequal conditions for businesses. While this may well sometimes be the case, the existing case study evidence unambiguously denies that this is a problem with a high practical relevance (Davidson 2006; Falkner et al. 2005; Jans et al. 2009; Morris 2011; Voermans 2009; see Thomann and Zhelyazkova (2017) for a wrap-up). My results support that these are rare, though real, situations. Beyond this, the present study shows that customization can have an important role for policy success. In particular, the complex configurations that facilitate successful implementation very often entail extensively customized EU policies. Conversely, an (almost) literal transposition of EU food safety policies is a prominent part of the story explaining unsuccessful implementation.

These findings also have important implications for policy implementation research more widely. On the one hand, results do suggest a link between policy outputs on paper and their outcomes in practice, once an evaluative perspective taking into account features of the policy problem, the implementing structures, the policy design and the political climate is adopted (Sabatier and Mazmanian 1980; Knill et al. 2012). On the other hand, they show that simple top-down or bottom-up views on discretion in implementation fall short of capturing the reality of policy implementation (Shapiro 1999). In tendency, however, the results tend to lend support to a hybrid bottom-up view which grants discretion, under certain circumstances, an important

role for policy success (Elmore 1979; Thomann et al., forthcoming).

It has to be noted that the results have limitations. First of all, customization does not equal compliance. Nothing in this study is to suggest that we should stop worrying about non-compliance. Rather, it is to suggest that we should move beyond a narrow focus on questions of compliance (Schmidt 2008). It is time to connect the dots: we know that most of customization happens within the boundaries of what EU rules allow for (Thomann and Zhelyazkova 2017), and that legal compliance can be decoupled from practical application (Zhelyazkova et al. 2016). The present results suggest that customization patterns can be an important missing part of the picture linking implementation on paper to implementation in practice (Versluis 2003, 2007). However, this evidence is limited to one regulatory policy in relatively similar, "old" EU member states. Future research will have to reveal whether a similar link exists in other countries and policy sectors.

Policy implementation in the EU should be evaluated according to empirical, rather than normative, criteria. The evidence suggests that member states and local policy implementers have a crucial role as problem-solvers when implementing EU rules. Using their discretion can help them to reach better outcomes in practice. The crucial question hence does not seem to be to avoid over-implementation at all costs, but to identify the conditions under which it can be reconciled with Europeanization dynamics, and even unfold a synergetic role.

Acknowledgements

References

Adam, C., Hurka, S. and Knill, C. (2015) 'Four styles of regulation and their implications for comparative policy analysis', *Journal of Comparative Policy Analysis: Research and Practice*, DOI:10.1080/13876988.2015.1082262.

Barrett, S. and Fudge, C. (1981) Policy and Action: Essays in the Implementation of Public Policy. London: Methuen.

Bauer, M.W. and C. Knill. 2014. A conceptual framework for the comparative analysis of policy change: Measurement, explanation and strategies of policy dismantling. *Journal of Comparative Policy Analysis: Research and Practice* 16(1): 28-44.

Baumgartner, Michael, and Alrik Thiem. 2015. Model Ambiguities in Configurational Comparative Research. *Sociological Methods & Research*.

Beugelsdijk, M. und S.CW. Eijffinger. 2005. The Effectiveness of Structural Policy in the European Union: An Empirical Analysis for the EU-15 in 1995–2001. *JCMS: Journal of Common Market Studies* 43(1): 37-51.

Bondarouk, E. and Liefferink, D. (2016) 'Diversity in sub-national EU implementation: the application of the EU Ambient Air Quality directive in 13 municipalities in the Netherlands', *Journal of Environmental Policy & Planning*, DOI:10.1080/1523908X.2016.1267612.

Bondarouk, E. and Mastenbroek, E. (2017, forthcoming). 'Reconsidering EU compliance: Implementation performance in the field of environmental policy', *Environmental Policy and Governance*, online first.

Bugdahn, S. 2005. Of Europeanization and domestication: the implementation of the environmental information directive in Ireland, Great Britain and Germany. *Journal of European Public Policy* 12(1): 177-199.

Werner Bussmann, Ulrich Klöti und Peter Knoepfel (Hrsg.). *Einführung in die Politikevaluation*. Basel, Frankfurt a.M.: Helbing & Lichtenhahn

Davidson, N. (2006) '*Davidson review: Implementation of EU legislation*', Retrieved online from http://www.cabinetoffice.gov.uk/regulation/reviewing_regulation/davidson_review/index.asp [last accessed 7.3.2016].

Dimitrova, Antoaneta, and Bernard Steunenberg. "Living in parallel universes? Implementing European movable cultural heritage policy in Bulgaria." *JCMS: Journal of Common Market Studies* 51.2 (2013): 246-263.

Dörrenbächer, N. 2017, forthcoming. Europe at the Frontline: Analyzing street-level motivations for the use of European Union migration law. *Journal of European Public Policy*.

Dörrenbächer, N. and Mastenbroek, E. (2017, forthcoming) 'Passing the Buck? Analysing the delegation of discretion after transposition of European Union law', *Regulation & Governance*, doi: 10.1111/rego.12153.

Dușa, A. 2007. "User Manual for the QCA(GUI) Package in R." *Journal of Business Research* 60(5):576-86.

Elmore, R.F. (1979) 'Backward mapping: Implementation research and policy decisions', *Political Science Quarterly* 94(4): 601-616.

Exadaktylos, T. & Zahariadis, N. (2014). Quid pro Quo: Political Trust and Policy Implementation in Greece during the Age of Austerity. *Politics & Policy* 42(1): 160–183.

Falkner, G., Treib, O., Hartlapp, M. and S. Leiber. 2005. *Complying with Europe: EU Harmonisation and Soft Law in the Member States.* New York: Cambridge University Press.

Gollata, J. A. M. and Newig, J. (2017) 'Policy implementation through multi-level governance: analysing practical implementation of EU air quality directives in Germany', *Journal of European Public Policy*, DOI: 10.1080/13501763.2017.1314539.

Gulbrandsen, C. 2011. The EU and the implementation of international law: the case of 'sea-level bureaucrats'. *Journal of European Public Policy* 18(7): 1034-1051.

Hartlapp, M. und G. Falkner. 2009. Problems of operationalization and data in EU compliance research. *European Union Politics* 10(2): 281–304.

Héritier, A. (1999) *Policy-making and diversity in Europe: Escape from deadlock*. Cambridge: Cambridge University Press.

Héritier, A. (2016) 'Rigour versus Relevance'? Methodological discussions in political science', *Politische Vierteljahresschrift* 57(1): 11-26.

Hinterleitner, M., Sager, F. and E. Thomann. 2016. The Politics of External Approval: Explaining the IMF's Evaluation of Austerity Programs. *European Journal of Political Research* 55(3): 549–567.

Hooghe, L. and Marks, G. (2003) 'Unraveling the central state, but how? Types of multilevel governance', American Political Science Review 97(2): 233–243.

Hooghe, Liesbet, Gary N. Marks, and Arjan H. Schakel. *The rise of regional authority: A comparative study of 42 democracies*. Routledge, 2010.

Howlett, Michael. 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, Michael, and Jeremy Rayner. 2007. "Design principles for policy mixes: cohesion and coherence in 'new governance arrangements'". *Policy and Society* 26(4): 1-18.

Jans, J.H., Squintani, L., Aragão, A., Macrory, R. and Wegener, B.W. (2009) '"Gold plating" of European environmental measures?, *Journal of European Environmental Planning Law* 6(4): 417- 435.

Jensen, C.B. 2007. Implementing Europe: A Question of Oversight. *European Union Politics* 8(4): 451-477.

Keman, H. (2000) 'Federalism and Policy Performance. A Conceptual and Empirical Inquiry', In Wachendorfer-Schmidt, U. *Federalism and Political Performance.* London: Routledge, pp. 196-227.

Knill, C. 1998. European policies: the impact of national administrative traditions. *Journal of Public Policy* 18(1): 1-28.

Knill, C. 2015. Implementation. In: Richardson, J. und S. Mazey. *European Union: power and policy-making*. London und New York: Routledge, 371-397.

Knill, C. and D. Lehmkuhl. 2002. The national impact of European Union regulatory policy: Three Europeanization mechanisms. *European Journal of Political Research* 41(2): 255-280.

Knill, C., Schulze, K. and J. Tosun. 2012. Regulatory Policy Outputs and Impacts: Exploring a Complex Relationship. *Regulation and Governance* 6(4): 427-444.

Knill, C. and J. Tosun. 2012. Governance Institutions and Policy Implementation in the European Union.

In Richardson, J. *Constructing a policy-making state: Policy dynamics in the EU*. Oxford: Oxford University Press, 309-333.

Knill, C., Tosun, J. and M.W. Bauer. 2009. Neglected faces of Europeanization: The differential impact of the EU on the dismantling and expansion of domestic policies. *Public Administration* 87(3): 519-537.

Knoepfel, P., Larrue, C., Varone, F. and Hill, M. (2011) *Public Policy Analysis*. Policy Press.

Mahoney, James, and Gary Goertz. "The possibility principle: Choosing negative cases in comparative research." *American political science review* 98.04 (2004): 653-669.

Majone, G. 1999. Regulation in comparative perspective. *Journal of Comparative Policy Analysis: Research and Practice* 1(3): 309-324.

Marsh, David, and Allan McConnell. "Towards a framework for establishing policy success." *Public administration* 88.2 (2010): 564-583.

Mastenbroek, E., van Voorst, S. and A. Meuwese. 2016. Closing the regulatory cycle? A meta evaluation of ex-post legislative evaluations by the European Commission. *Journal of European Public Policy* 23(9): 1329-1348.

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.

McCubbins, Mathew D. and Schwartz, Thomas, 1984, "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms", American Journal of Political Science, 28(1): 165–179,

Medzihorsky, J., Oana, I., Quaranta, M. and C.Q. Schneider. 2017. *SetMethods: A Package Companion to "Set-Theoretic Methods for the Social Sciences". R* Package Version 2.1. URL: http://cran.r-project.org/package=SetMethods.

Morris, R.KA. (2011) 'The application of the Habitats Directive in the UK: compliance or gold plating?', *Land Use Policy* 28(1): 361-369.

Patton, Michael Quinn (1997). *Utilization-focused Evaluation*. 3rd ed. Thousand Oaks: Sage.

Ragin, C.C. (2009). "Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA)". In Rihoux, B. and C. C. Ragin. *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. Los Angeles, London, New Delhi and Singapore: Sage Publications, 87-121.

Rihoux, B., and C.C. Ragin. 2009. *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks and London: Sage.

Rutz, S., Mathew, D., Robben, P. and A. Bont. 2015. Enhancing responsiveness and consistency: Comparing the collective use of discretion and discretionary room at inspectorates in England and the Netherlands. *Regulation & Governance*. DOI:10.1111/rego.12101.

Sabatier, P. and Mazmanian, D. (1980) 'The Implementation of Public Policy: A Framework of Analysis', *Policy Studies Journal* 8(4): 538-560.

Sager, Fritz, and Christian Rüefli. "Die Evaluation öffentlicher Politiken mit föderalistischen Vollzugsarrangements. Eine konzeptionelle Erweiterung des Stufenmodells und eine praktische Anwendung." *Swiss Political Science Review* 11.2 (2005): 101-129.

Sager, F., Thomann, E., Zollinger, C. and C. Mavrot. 2014. Confronting theories of European integration: A comparative congruence analysis of veterinary drugs regulations in five countries. *Journal of Comparative Policy Analysis: Research and Practice* 16(5): 457-474.

Sager, F., Thomann, E., Zollinger, C. and C. Mavrot. 2011. *Tierarzneimittelregulierung in Europa*.

Study mandated by the Swiss Federal Office of Public Health. Bern: Center of Competence for Public Management.

Schaffrin, A., Sewerin, S. and S. Seubert. 2015. Toward a Comparative Measure of Climate Policy Output. *Policy Studies Journal* 43(2): 257-282.

Scharpf, F. (1997) 'Introduction: the problem-solving capacity of multilevel governance', *Journal of European Public Policy* 4(4): 520-538.

Schneider, C.Q., and C. Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis.* New York: Cambridge University Press.

Shapiro, M. (1999) 'Implementation, discretion and rules', in J.A.E. Vervaele (ed.), *Compliance and enforcement of European Community law,* The Hague: Kluwer Law International, pp.27-34.

Skjærseth, J.B. und J. Wettestad. 2008. Implementing EU emissions trading: success or failure? *International Environmental Agreements: Politics, Law and Economics* 8(3): 275-290.

Spendzharova, Aneta, and Esther Versluis. "Issue salience in the European Policy Process: what impact on transposition?." *Journal of European Public Policy* 20.10 (2013): 1499-1516.

Thomann, E. 2015a. Customizing Europe: Transposition as bottom-up implementation. *Journal of European Public Policy* 22(10): 1368-1387.

Thomann, E. (2015b) 'Is output performance all about the resources? A fuzzy-set Qualitative Comparative Analysis of street-level bureaucrats in Switzerland', *Public Administration* 93(1): 177-194.

Thomann, E., Lieberherr, E. and K. Ingold. 2016. Torn between state and market: Private policy implementation and conflicting institutional logics. *Policy & Society* 35(1): 57-69.

Thomann, E. and M. Maggetti. 2017, forthcoming. Designing research with Qualitative Comparative Analysis (QCA): Approaches, challenges, tools. Conditionally accepted for publication in *Sociological Methods & Research.*

Thomann, E. and F. Sager. 2017a. Moving beyond legal compliance: Innovative approaches to EU multi-level implementation. *Journal of European Public Policy*, DOI:10.1080/13501763.2017.1314541.

Thomann, E. and F. Sager. 2017b. Toward a better understanding of implementation performance in the EU multi-level system. *Journal of European Public Policy*, DOI: 10.1080/13501763.2017.1314542.

Thomann, E. and F. Sager. 2017c. Hybridity in action: Accountability dilemmas of public and for-profit food safety inspectors in Switzerland. In: Paul Verbruggen and Tetty Havinga (Eds). *Hybridization of food governance: Trends, types and results*. Cheltenham and Massachusetts: Edward Elgar, 100-120.

Thomann, E., van Engen, N. and L. Tummers. Under review. The necessity of discretion: a behavioral evaluation of bottom-up implementation theory.

Thomann, E. and A. Zhelyazkova. 2017. Moving beyond (non-)compliance: the customization of European Union policies in 27 countries. *Journal of European Public Policy*, DOI: 10.1080/13501763.2017.1314536.

Toshkov, D. und L. de Haan. 2013. The Europeanization of asylum policy: an assessment of the EU impact on asylum applications and recognitions rates. *Journal of European Public Policy* 20(5): 661-683.

Toshkov, Dimiter. "Compliance with EU Law in Central and Eastern Europe." *L'Europe en Formation* 2 (2012): 91-109.

Toshkov, D. (2016) *Research Design in Political Science*. London and New York: Palgrave Macmillan.

Tosun, J. 2012. Environmental Monitoring and Enforcement in Europe: A Review of Empirical Research. *Environmental Policy and Governance* 22: 437–448.

Treib, O. 2014. Implementing and complying with EU governance outputs. *Living Reviews in European Governance* 9(1).

Tummers, Lars, and Victor Bekkers. "Policy implementation, street-level bureaucracy, and the importance of discretion." *Public Management Review* 16.4 (2014): 527-547.

Versluis, E. 2003. Enforcement matters. Enforcement and Compliance of European Directives in Four Member States. Delft: Eburon.

Versluis, E. 2007. Even rules, uneven practices: Opening the 'black box' of EU law in action. *West European Politics* 30(1): 50-67.

Voermans, W. 2009. Gold-plating and double banking: an overrated problem? In Snijders, H. and Vogenauer, S. *Content and Meaning of National Law in the Context of Transnational Law*. Munich: Sellier European Law Publishers, 79-88.

Whitford, A. (2007) 'Decentralized Policy Implementation', *Political Research Quarterly* 60(1): 17-30.

Winter, S.C. (2003) ' Implementation Perspectives: Status and Reconsideration', In: B.G. Peters and J. Pierre, eds. *Handbook of Public Administration*. Thousand Oaks: Sage, pp. 212-222.

Zhelyazkova, A, Kaya, C. and Schrama, R. (2016) 'Decoupling practical and legal compliance: Analysis of member states' implementation of EU policy', *European Journal of Political Research* 55(4): 827-846.

Appendix

Appendix A: Supplementary information about sources and cases

Table A1: Policies, EU legal basis, target groups, indicators for outputs and outcomes

	Policy id	Policy content	Addressees	EU directive	Outputs (practical transposition)	Outcomes (behavioral change)
	D1	Requirement of clinical examination of animals prior to prescribing drug	Veterinarians Livestock holders	90/167/EEC[4]	Controls on farms to monitor compliance	Veterinarians prescribe drugs to animals with due knowledge of their health status
	D2	Actors authorized to dispense prescription drugs	Veterinarians Pharmacies	2001/82/EC[5]	Controls in dispensaries to monitor compliance If applicable: training opportunities are provided Authorizations are issued properly	Prescription drugs are only dispensed by competent actors
	D3	Actors authorized to dispense medicated feed	Veterinarians Feed mills	90/167/EEC	Controls in dispensaries to monitor compliance If applicable: training opportunities are provided Authorizations are issued properly	Medicated feed are only dispensed by competent actors
	D4	Amount of prescription drugs which may be dispensed	Veterinarians Livestock holders	2001/82/EC	Controls in dispensaries and on farms to monitor compliance	Safe & moderate use of prescription drugs
	D5	Amount of medicated feed which may be dispensed	Veterinarians Livestock holders	90/167/EEC	Controls in dispensaries and on farms to monitor compliance	Safe & moderate use of medicated feed
	D6	Dispensing categories	Veterinarians Pharmacies	2001/82/EC	Products are classified accordingly and information is made available to users Inspectors refer to these categories in controls	Drugs are made available to end users with appropriate consideration of risks
Dispensing	D7	Dispensing rights of veterinarians and pharmacies exceeding the mere	Veterinarians Pharmacies	2001/82/EC	Controls in dispensaries to monitor compliance	Vets & pharmacies produce and repack drugs under safe and controlled conditions

47

	Policy id	Policy content	Addressees	EU directive	Outputs (practical transposition)	Outcomes (behavioral change)
		distribution of drugs				
	D8	Required duration of storage of dispensing documentation, by whom	Veterinarians Pharmacies Livestock holders	2001/82/EC	Controls in dispensaries and on farms to monitor compliance	Traceability of dispensing documentation for a sufficient time
	D9	Required duration of storage of prescription and by whom	Veterinarians Livestock holders	2001/82/EC	Controls in dispensaries and on farms to monitor compliance	Traceability of prescription for a sufficient time
	D10	Actors authorized to manufacture drugs which do not require a market authorization	Veterinarians Pharmacies Livestock holders	2001/82/EC	Controls in dispensaries to monitor compliance	Veterinary drugs are manufactured for commercial or industrial use only by competent actors
	D11	Medicated feed requires a prescription on a standardized form	Veterinarians Feed mills Livestock holders	90/167/EEC	Prescription templates are provided and mandatory	Medicated feed is given to end users under uniform, controlled conditions in EU single market
	D12	Actors authorized to prescribe veterinary drugs	Veterinarians	2001/82/EC	Controls in dispensaries to monitor compliance If applicable: training opportunities are provided	Veterinary drugs are only prescribed by competent actors
	D13	Exemptions from prescription requirement	Veterinarians Pharmacies	2006/130/EC[6]	Products are classified accordingly	Drugs that do not pose any risk even if used incorrectly are available without prescription
Administration	A1	Actors who may administer the drug used off-label	Veterinarians Livestock holders	2001/82/EC	Controls on farms to monitor compliance If applicable: training opportunities are provided	Off-label use is only performed by competent persons
	A2	Possibilities for off-label use in cases of supply shortage	Veterinarians	2001/82/EC	Controls on farms to monitor compliance	Safe off-label use in line with EU cascade rule Adequate treatment in case of supply shortage
	A3	Required duration of storage of administration documentation, by whom	Veterinarians Livestock holders	2001/82/EC	Controls on farms to monitor compliance	Traceability of administration documentation for a sufficient time
	A4	Possibility of on-farm manufacturing of medicated	Livestock holders	90/167/EEC	Controls on farms to monitor compliance	Livestock owners manufacture medicated feed on farm only if

Policy id	Policy content	Addressees	EU directive	Outputs (practical transposition)	Outcomes (behavioral change)
	feed for livestock owners			If applicable: training opportunities are provided	competent and with appropriate machinery
A5	Is top dressing (manual adding of drug into feed) allowed?	Livestock holders	90/167/EEC	Authorities issue information / guidelines Controls on farms to monitor compliance	Top dressing is done safely in accordance with summary of product characteristics
A6	Withdrawal periods	Veterinarians Livestock holders	2001/82/EC	Products are classified accordingly Controls in slaughterhouses and supermarkets to monitor compliance	Animals are not processed into products that enter the market while there could still be residues of drugs

Table A2: Types and frequency of reported implementation problems (multiple possible)

Outputs	Frequency		Outcomes	Frequency	
Regional differences in enforcement	AU	0	Rule hinders target groups from contributing to policy goal	AU	1
	CH	10		CH	3
	FR	0		FR	1
	GE	0		GE	1
	UK	10		UK	1
Sectoral differences in enforcement	AU	0	Rule ineffective in practice (e.g. incentive structures, lacking regulation)	AU	0
	CH	0		CH	0
	FR	0		FR	3
	GE	7		GE	0
	UK	0		UK	1
Enforcement deficit (ineffective or insufficient)	AU	0	Non-compliance with rule	AU	1
	CH	11		CH	3
	FR	9		FR	3
	GE	7		GE	1
	UK	10		UK	1
Monitoring deficit	AU	0			
	CH	7			
	FR	0			
	GE	6			
	UK	0			
Insufficient implementation orders / information directed toward addressees (incomplete, inaccurate, dysfunctional)	AU	0			
	CH	2			
	FR	2			
	GE	0			
	UK	1			
At least 1 problem	**AU**	**0**		**AU**	**2**
	CH	**18**		**CH**	**6**
	FR	**11**		**FR**	**6**
	GE	**8**		**GE**	**2**
	UK	**11**		**UK**	**2**

N/country = 19.

Table A3: FVO reports included in analysis

Food and Veterinary Office (FVO). 2008. REPORT OF A SPECIFIC AUDIT CARRIED OUT IN GERMANY FROM 22 TO 29 SEPTEMBER 2008 IN ORDER TO EVALUATE THE CONTROL OF RESIDUES AND CONTAMINANTS IN LIVE ANIMALS AND ANIMAL PRODUCTS, INCLUDING CONTROLS ON VETERINARY MEDICINAL PRODUCTS. PART B – SECTOR SPECIFIC ISSUES. DG(SANCO)/2008/7775 MR – Final.
Food and Veterinary Office (FVO). 2010. FINAL REPORT OF A SPECIFIC AUDIT CARRIED OUT IN FRANCE FROM 22 FEBRUARY TO 01 MARCH 2010 IN ORDER TO EVALUATE THE CONTROL OF RESIDUES AND CONTAMINANTS AND THE USE OF VETERINARY MEDICINAL PRODUCTS IN FOOD PRODUCING ANIMALS IN THE CONTEXT OF A GENERAL AUDIT. DG(SANCO) 2010-8435 - MR FINAL.
Food and Veterinary Office (FVO). 2011. FINAL REPORT OF A MISSION CARRIED OUT IN SWITZERLAND FROM 17 TO 21 JANUARY 2011 IN ORDER TO EVALUATE THE MONITORING OF RESIDUES AND CONTAMINANTS IN LIVE ANIMALS AND ANIMAL PRODUCTS, INCLUDING CONTROLS ON VETERINARY MEDICINAL PRODUCTS. DG(SANCO) 2011-8907 - MR FINAL.
Food and Veterinary Office (FVO). 2011. FINAL REPORT OF AN AUDIT CARRIED OUT IN AUSTRIA FROM 14 TO 20 JUNE 2011 IN ORDER TO EVALUATE THE MONITORING OF RESIDUES AND CONTAMINANTS IN LIVE ANIMALS AND ANIMAL PRODUCTS, INCLUDING CONTROLS ON VETERINARY MEDICINAL PRODUCTS. DG(SANCO) 2011-8910 - MR FINAL.
Food and Veterinary Office (FVO). 2009. FINAL REPORT OF A SPECIFIC AUDIT CARRIED OUT IN THE UNITED KINGDOM FROM 16 FEBRUARY TO 23 FEBRUARY 2009 IN ORDER TO EVALUATE THE CONTROL OF RESIDUES AND CONTAMINANTS AND THE USE OF VETERINARY MEDICINAL PRODUCTS IN FOOD PRODUCING ANIMALS IN THE CONTEXT OF A GENERAL AUDIT. DG(SANCO)/ 2009-8128 - MR - FINAL.

Table A4: Selection criteria for non-EU member countries

European non-EU member country	Democratically stable and peaceful since Maastricht treaty in 1993[1]	Value added to GDP by agriculture in per cent s within the range of 1 (UK) and 4.7 (France) [2]	Trade surplus for fishery products[3]
Switzerland	Yes	Yes (1.6)	No
Norway	Yes	Yes (2.1)	Yes
Moldova	Relatively	No (29)	
Albania	No		
Belarus	No		
Bosnia	No		
Croatia	No		
Kosovo	No		
Macedonia	No		
Montenegro	No		
Serbia	No		
Ukraine	No		

Notes: GDP = Gross domestic product.   EU = European Union.   UK = United Kingdom.

European countries according to United Nations with non-EU membership by 2010. Does not include the very weakly populated countries of Iceland and Liechtenstein.

[1]Not given if periods of non-democratically elected leadership or significant internal or external conflicts involving armed violence have occurred.

[2]Sources: Switzerland: http://de.statista.com. Other countries: World Development Indicators. Reference year: 2000.

[3]Source: Switzerland: Federal Statistical Office. Norway: Eurostat. Fishery statistics. Data 1990-2006. Reference year: 2000.

Appendix B: Discussion of measurement and calibration

Table B1 displays the descriptive statistics of the raw variables and the set skewness.

Table B1: Descriptive statistics of raw variables, set skewness

Variable	Min-imum	Max-imum	Mean	Median	Standard devia-tion	Set skewness
Successful implementation	0	1	0.57	0.7	0.34	62.1
Customized density	0	2	0.89	1	0.81	-
Customized restrictiveness	0	2	0.94	1	0.91	-
Customization	0	4	1.92	2	1.26	70.5
Problem tractability	2	6	3.89	4	1.30	57.9
Regional self-rule	7.75	15	10.95	12	2.75	60
Integrated implementation structure	8	11	9.60	10	1.03	40
Centralized implementation structure						60
Active enforcement system	0	1	0.72	1	0.38	67.4
External coherence	0	1	0.60	1	0.49	-
Suitable enforcement structure	0	1	0.80	1	0.40	-
Coherent problem definition	0	1	0.40	0	0.49	-
Coherent objectives	0	1	0.60	1	0.49	-
Coherent functional parameters	0	1	0.40	0	0.49	-
Coherent organizational parameters	0	1	0.40	0	0.49	-
Policy design coherence	2	4	3.20	4	0.98	60
Domestic resistance	2	6	3.81	3	1.48	36.8

N = 95.
Skewness= per cent of cases with set membership > 0.5.

Successful implementation. On the basis of a multi-annual programme, each year DG Health and Food Safety produces a work programme in consultation with other Commission services and with Member States, which considers risk and trade factors, plus the status of legislation, to prioritize visits. Drawing from a team of some 170 professionals from most EU Member States, audit teams are typically composed of two auditors, often with the presence of a national expert from a Member State authority. Methods include on-the-spot audits, preceded by desk based exercises and collation of Member States data. The audits focus on the control system

rather than individual premises, typically entailing visiting the control authority, a number of regional and local authorities, laboratories and a number of accompanied site visits (e.g. to farms, processors, feed units, slaughterhouses and retailers). Using several sources is useful to enhance the robustness of and confidence in the results and to cross-validate the coding resulting from the FVO reports. While Sager et al. (2011) did not evaluate the implementation of the EU rules in the five countries, they did systematically report on implementation problems, based on a targeted analysis of national policy documents and expert interviews with public administrators and representatives of target group organizations.

Using these data sources has advantages and disadvantages. The advantages are considerable: I have directly comparable data for all five countries, covering the exact same policies, using identical methodologies for assessing policy outcomes[12] and being reported in very similar structures. A disadvantage is that the FVO did not collect the data at identical points in time for the five countries (Austria: 2011, France: 2010, Germany: 2008, UK: 2009, Switzerland: 2011).

My coding strategy proceeds in three steps: In a first step, I identify for each of the 16 EU rules the indicators for a) outputs (practical transposition, including monitoring and enforcement) and b) outcomes (behavioral change of addressees) that are explicitly available in one or both of the two datasources (FVO reports and Sager et al. 2011), see Table A1. I considered availability to be given if the output or outcome was mentioned in one of the two sources for at least one country. In a second step, I coded these outputs and outcomes for each country, first based on the FVO reports and then using Sager et al. (2011). The guiding question was: are problems reported relating to this specific output or outcome? If the output or outcome was specifically mentioned in a report (positively or negatively), the content was documented. In a final step, I performed additional desk-based research on those cases where the available

---

[12] See http://ec.europa.eu/food/audits_analysis_en for a more detailed description (last accessed 12.5.2017).

information was ambiguous, or my two data sources suggested contradictory results. As discussed earlier, I aggregated the values for outputs and outcomes into a 6-value fuzzy set (Table 2; indirect calibration).

Problem tractability. Minor issues (coded 0) are EU rules that clearly concern minor issues representing refinements. They refer to very rare situations, merely administrative procedures, and / or imply only negligible costs or benefits for the addressees. Macro-issues (coded 3) are then issues that refer to frequently occurring situations and have notable consequences for the addressees. On the other hand, I count the number of principal target groups for each rule.[13] The added index of problem tractability ranges from 1 to 6. Problem tractability is (rather) high for minor issues, as well as for major issues with only one target group (four-value fuzzy set; indirect method of calibration),

Centralized implementation structure. First, to measure regional self-rule in a country (REG), I use Hooghe et al.'s (2010) index of regional self-rule in 2006, which entails the dimensions of institutional depth, policy scope, fiscal autonomy, and representation, and ranges from 0 to 15.[14] Self-rule is more given than absent if the relevant regions in a country achieve at least two thirds of the possible scores (10 or more; direct method of calibration). Second, I measure the degree of integration of the implementation structure (INT) as the number of institutions involved in transposing, monitoring and enforcing the respective rule for a given unit (e.g., a veterinary dispensary or a farmer) at different organizational tiers.[15] A four-value fuzzy set captures the diversity of the five countries (indirect calibration method). I consider an

---

[13] The coding of the number of target groups refers to the EU rule. Note that in a few isolated cases, the respective rule may target more actors in a specific country. For example, the regulation of the on-farm manufacture of medicated feed (OFM) targets livestock owners. However, in Switzerland, the rules for OFM also involve duties for veterinarians. The analysis captures this through customized restrictiveness.

[14] France: Départements, Germany: Länder, Switzerland: cantons, UK: average value for Grater London, Northern Ireland, Scotland and Wales. England, Scotland, Wales and Northern Ireland each have their own government or executive.

[15] Source: Sager et al. (2011: 353). In the UK, the situation in England serves as a reference.

implementation structure to be centralized in a country for a given rule, if regions have a low degree of self-rule, or the number of implementing units is comparatively low. One factor can substitute for the other: even if there are many institutions involved, in a political system with low regional authority, these bodies tend to be tightly steered from above. Similarly, even if regions have strong authority, that may not prove problematic in the specific policy sector if the need for coordinating bodies is low.

Active enforcement system. In some countries, enforcement strategies vary for policies on veterinary drugs dispensing and veterinary drugs administration. Sometimes, the enforcement strategy on paper differs from how it is practiced (e.g., due to the high amount of livestock farms, or lacking resources) (Thomann 2015b). Finally, some countries have implemented hybrid enforcement structures which involves private enforcement agents (e.g., veterinarians; in force in Austria and Switzerland), where public and private enforcement strategies can differ (Thomann and Sager 2017c). I thus assess the enforcement strategy separately for dispensing and administration rules. In my cases, the enforcement strategy is fully passive (0), partly passive (0.33; legally active, but de facto passive), partly active (0.67; public passive, private active), or fully active (1) (indirect calibration).

Coherent policy design (COH).

Coherence may be fully given (1) or partly or fully absent (0) for six aspects: external coherence; suitability of enforcement structures for service provision; internal coherence of problem definition, objectives, functional, and organizational parameters. Functional parameters define the operational measures that are necessary to achieve the policy objectives. They determine immediate rights and duties of the addressees, thereby also influencing the extent and quality of the services delivered. Operational parameters refer to the executive agencies responsible for different aspects of implementation (Bussmann et al. 1997: 87-88; Knöpfel et al. 2011: 169-174; 219). The resulting added index ranges from 0 to 6. I consider a

policy design as more coherent than incoherent if the majority of these elements are considered as fully coherent by legal experts and target groups (see Sager et al. 2011). In the given dataset, I only observe values of 2 and 4. To a avoid distorted parameters of fit due to inevitably less-than-perfect set memberships in truth table rows, the final set COH is coded as a dichotomous crisp set.

Strong domestic resistance (RES). I construct an added domestic resistance index. The strength of domestic resistance hinges on the power of the opposed stakeholder groups to successfully influence policy making, which was evaluated by the interviewees for each country (absent (1), medium (2) or significant (3)). The final condition 'domestic resistance' is an added index: it is fully given if one of the target group opposes the policy (3) and that group is powerful (sum: 6); and more given than not if a target group opposes the policy, but is only moderately or not influential (sum: 4 or 5). If no one opposes the policy (0), but one of the policy's target groups is influential, then resistance is mostly (3), and if they are all not or only moderately influential, then fully absent (1 or 2) (indirect method of calibration). For policies with several addressees, the most powerful target group served as point of reference.

Figure B1: Raw values and calibrated fuzzy values with crossover points
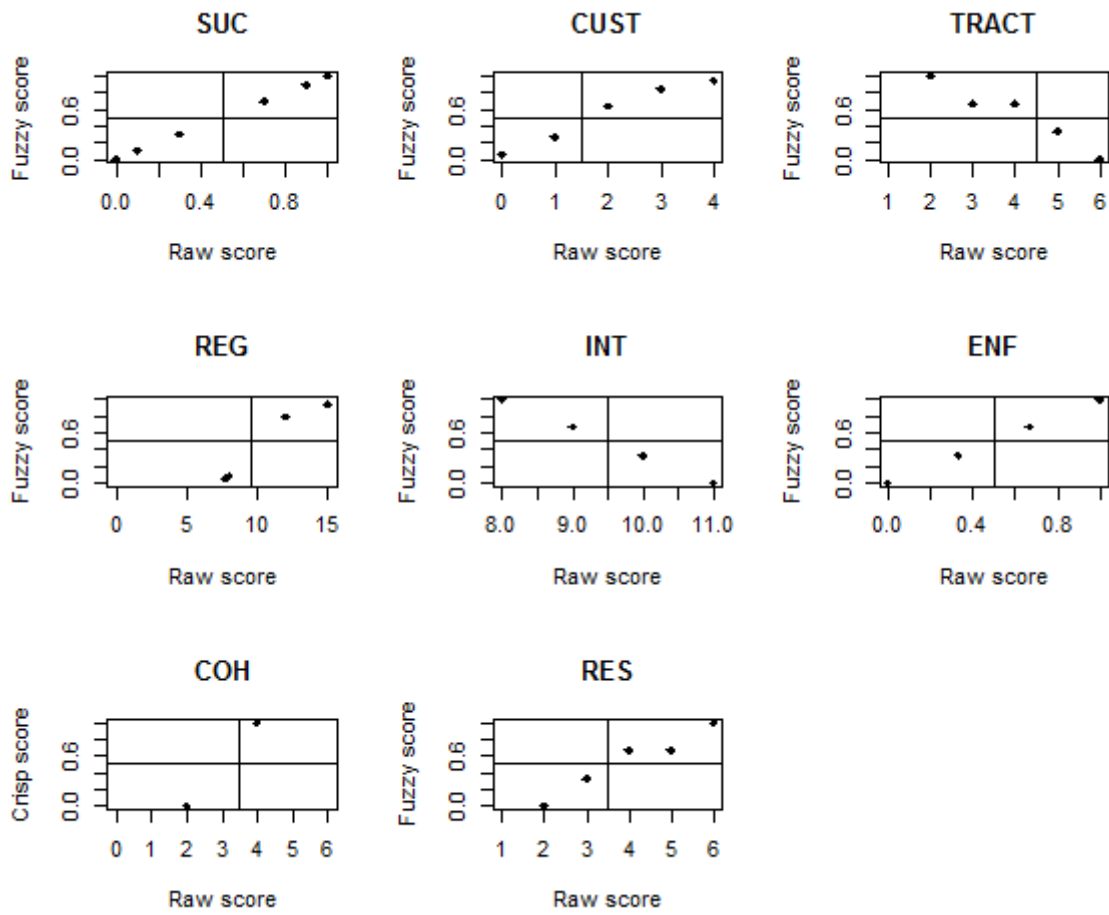
Appendix C: Documentation of analyses of necessity and sufficiency

Table C1: Analysis of necessity

	Successful implementation SUC			Unsuccessful implementation ~SUC		
Condition	Consistency	RoN	Coverage	Consistency	RoN	Coverage
CUST	0.703	0.717	0.702	0.658	0.601	0.498
TRACT	0.713	0.679	0.68	0.721	0.586	0.522
CENT	0.731	0.597	0.638	0.785	0.527	0.52
ENF	0.727	0.479	0.575	0.808	0.432	0.485
COH	0.609	0.612	0.577	0.588	0.536	0.423
RES	0.58	0.766	0.677	0.618	0.7	0.548
~CUST	0.497	0.794	0.657	0.606	0.771	0.608
~TRACT	0.498	0.832	0.701	0.557	0.785	0.596
~CENT	0.45	0.875	0.734	0.453	0.81	0.561
~ENF	0.349	0.897	0.705	0.292	0.822	0.448
~COH	0.391	0.771	0.555	0.412	0.73	0.445
~RES	0.613	0.747	0.679	0.636	0.671	0.535

Notes: To count as necessary, a condition has to meet the following thresholds: consistency 0.9, coverage 0.6, RoN 0.5. No complex union of sets passes these thresholds.

Box C1: Setting the raw consistency threshold

Measurement error and discretionary analytic decisions can affect the robustness of QCA results (Maggetti and Levi-Faur 2013; Thiem et al. 2015). In particular, how raw consistency thresholds are set decisively affects the results. Different parameters help researchers decide about appropriate raw consistency thresholds: Raw consistency values indicate the sufficiency of a row for the outcome, PRI values, its simultaneous sufficiency for the negated outcome, and deviant cases consistency in kind (DCCK) indicate cases that "falsify" the sufficiency claim (Schneider and Wagemann 2012). Especially with large-N analyses, the parameters of fit can sometimes be misleading (Haesebrouck 2015). In order to account for possible pitfalls, I use all these criteria for setting raw consistency thresholds, checking for different options to enhance the robustness of and confidence in my results.

Successful implementation (SUC)

As Table C2 shows, raw consistency values and PRI values tend to be generally high. One reason for this is that the outcome set is slightly skewed toward high membership. However, this makes it difficult to identify an obvious choice for a raw consistency threshold. Generally speaking, the parameters of fit are difficult to interpret in analysis using both a large N and

fuzzy sets, since they are heavily influenced by the scores of substantively irrelevant cases (Haesebrouck 2015). I therefore additionally use the presence of DCCK as an indicator.

Option 1. A first obvious decline in PRI values is found in row 64. Indeed, there we see that the share of DCCK is 50 per cent, whereas it is much lower in the higher-ranked rows. The subsequent row 27, too, has one-fourth DCCK. A first, obvious threshold is therefore 0.87 (including row 63, excluding row 64). The results have two weaknesses: first, they are ambiguous (2 parsimonious models) when including logical remainders (Baumgartner and Thiem 2015). Second, coverage is low with 0.553. Indeed, given that there are many more rows that are highly consistent, this threshold may appear overly strict.

Option 2. Therefore, I run a new analysis that includes three more rows with no DCCK (rows 17, 45 and 2), while excluding all rows with 25% DCCK or more. Accordingly, I apply a raw consistency threshold of 0.816 and exclude rows 48, 59, 32, 27, 64 and 15. This solution, however, produces more complex results and more model ambiguity when including logical remainders (4 parsimonious models).

Option 3. As a third option, I include two more rows (49 and 27) which each have only one DCCK (25%). This will increase coverage, while it produces two more DCCK in the solution (total of 3) (Raw consistency threshold 0.806, excluding rows 64, 32, 59, 15, 48, 23 and 3). This option, again, produces four parsimonious models.

Option 4. Under this option, I use PRI values as a prime indicator, specifically a sharp decline in PRI values between rows 45 (0.680) and 59 (0.553). Accordingly, I set the threshold above row 59 (0.8315), and exclude rows 27 and 64 with 25% or more DCCK. This option produces the same levels of ambiguity as option 1, while covering more cases. Overall, option 4 is thus the most promising strategy for coding the truth table for SUC.

Unsuccessful implementation (~SUC)

In contrast to the truth table for CUST, raw consistencies and PRI values are generally low in the analysis for cust. Again, this may partly be due to the outcome being skewed toward low membership.

Option 1. A first option is to set the raw consistency threshold at the usually recommended lower limit of 0.75. This option, however, results in extremely low coverage; only 9 cases can be explained.

Option 2. Accordingly, in a second attempt I lower the raw consistency threshold to 0.7. This

threshold is lower than what is usually recommended; however, parameters of fit and therefore also thresholds are highly contingent on the specific research context (Schneider and Wagemann 2010). Here they seem to be artificially low as an artefact of calibration and case distributions (Thomann and Maggetti 2017); lowering the threshold seems appropriate to correct for this. To nonetheless facilitate an accurate analysis of sufficient conditions, I make sure that rows with 25% or more DCCK are excluded from the analysis (rows 64, 16). The resulting solutions have a consistency well above 0.75, a good coverage, and moderate levels of ambiguity (2 parsimonious models). Option 2 is therefore chosen for the analysis.

Table C2: Truth table for successful implementation (SUC)

Row Nr.	CUST	TRACT	CENT	ENF	COH	RES	OUT	n	Raw cons.	PRI	% DCCK
19	0	1	0	0	1	0	1	1	0.969	0.931	0
50	1	1	0	0	0	1	1	4	0.933	0.881	0
34	1	0	0	0	0	1	1	5	0.925	0.899	0
47	1	0	1	1	1	0	1	6	0.919	0.854	0
31	0	1	1	1	1	0	1	5	0.908	0.832	0
56	1	1	0	1	1	1	1	2	0.893	0.814	0
16	0	0	1	1	1	1	1	1	0.890	0.722	0
11	0	0	1	0	1	0	1	1	0.887	0.764	0
63	1	1	1	1	1	0	1	9	0.881	0.799	11
64	1	1	1	1	1	1	0	2	0.866	0.689	50
27	0	1	1	0	1	0	0	4	0.863	0.730	25
17	0	1	0	0	0	0	1	1	0.855	0.752	0
32	0	1	1	1	1	1	0	1	0.842	0.643	100
45	1	0	1	1	0	0	1	2	0.832	0.680	0
59	1	1	1	0	1	0	0	1	0.831	0.553	100
15	0	0	1	1	1	0	0	3	0.829	0.630	100
48	1	0	1	1	1	1	0	5	0.825	0.656	40
2	0	0	0	0	0	1	0	1	0.821	0.724	0
23	0	1	0	1	1	0	0	1	0.814	0.683	100
3	0	0	0	0	1	0	0	1	0.813	0.614	100
49	1	1	0	0	0	0	0	4	0.810	0.725	25
46	1	0	1	1	0	1	0	3	0.805	0.675	33
18	0	1	0	0	0	1	0	2	0.798	0.630	50
51	1	1	0	0	1	0	0	1	0.791	0.583	100
55	1	1	0	1	1	0	0	3	0.783	0.656	67
40	1	0	0	1	1	1	0	3	0.767	0.635	67
61	1	1	1	1	0	0	0	6	0.756	0.558	33
62	1	1	1	1	0	1	0	2	0.753	0.478	100
33	1	0	0	0	0	0	0	2	0.708	0.627	50
39	1	0	0	1	1	0	0	4	0.694	0.528	100
52	1	1	0	0	1	1	0	3	0.686	0.522	67
29	0	1	1	1	0	0	0	3	0.670	0.345	67
14	0	0	1	1	0	1	0	1	0.634	0.397	100
13	0	0	1	1	0	0	0	2	0.565	0.254	100

Limited diversity: 34 out of 64 truth table rows are observed empirically (53.1%). Limited diversity is due to clustered, not arithmetic remainders.

Raw consistency threshold 0.8315), rows 27, 32 and 64 are excluded. Directional expectations see Table 1.

Conservative solution:

CUST*cent*enf*coh*RES + cust*TRACT*cent*enf*res + CUST*tract*CENT*ENF*res + TRACT*CENT*ENF*COH*res + cust*tract*CENT*enf*COH*res + cust*tract*CENT*ENF*COH*RES + CUST*TRACT*cent*ENF*COH*RES → SUC (solution consistency 0.898, solution coverage 0.621).

Parsimonious solutions:

The present data display tied logically redundant prime implicants, and hence, a modest degree of model

62

ambiguity. All solutions are reported below (Baumgartner and Thiem 2015; Schneider and Wagemann, 2012: 108ff). The two parsimonious models differ only in path 7 regarding the role of cent (M1) versus enf (M2), respectively.

M1: tract*CENT*enf + cust*tract*COH*RES + TRACT*cent*ENF*RES + CUST*tract*CENT*res + TRACT*CENT*ENF*COH*res + cust*TRACT*cent*enf*res + (CUST*cent*coh*RES) → SUC (solution consistency 0.880, solution coverage 0.643).

M2: tract*CENT*enf + cust*tract*COH*RES + TRACT*cent*ENF*RES + CUST*tract*CENT*res + TRACT*CENT*ENF*COH*res + cust*TRACT*cent*enf*res + (CUST*enf*coh*RES) → SUC (solution consistency 0.880, solution coverage 0.643).

Intermediate solutions:

M1 (PS M1): CUST*cent*coh*RES + cust*tract*ENF*COH*RES + tract*CENT*enf*COH*res + TRACT*CENT*ENF*COH*res + cust*TRACT*cent*enf*res + CUST*tract*CENT*ENF*res + (CUST*TRACT*cent*ENF*RES)→ SUC (solution consistency 0.894, solution coverage 0.626).

M2 (PS M1): CUST*cent*coh*RES + cust*tract*ENF*COH*RES + tract*CENT*enf*COH*res + TRACT*CENT*ENF*COH*res + cust*TRACT*cent*enf*res + CUST*tract*CENT*ENF*res + (TRACT*cent*ENF*COH*RES) → SUC (solution consistency 0.894, solution coverage 0.626).

M3 (PS M2): CUST*enf*coh*RES + cust*tract*ENF*COH*RES + tract*CENT*enf*COH*res + TRACT*cent*ENF*COH*RES + TRACT*CENT*ENF*COH*res + cust*TRACT*cent*enf*res + CUST*tract*CENT*ENF*res → SUC (solution consistency 0.896, solution coverage 0.621).

Apart from the difference reflected in the two parsimonious models, the three intermediate models differ in the relevance they grant to CUST (M1) as opposed to COH (M2, M3) together with TRACT*cent*ENF*RES. I present M1 in the running text because it is most insightful regarding the role of CUST. The other two models suggest that a coherent policy design could be an equally plausible INUS condition in this configuration.

Table C3: Simplifying assumptions and easy counterfactual (SUC)

Row	CUST	TRACT	CENT	ENF	COH	RES	Easy counterfactual (used for intermediate solution)
4	0	0	0	0	1	1	
8	0	0	0	1	1	1	X
9	0	0	1	0	0	0	
10	0	0	1	0	0	1	
12	0	0	1	0	1	1	
22	0	1	0	1	0	1	
24	0	1	0	1	1	1	x
38	1	0	0	1	0	1	x
41	1	0	1	0	0	0	
42	1	0	1	0	0	1	
43	1	0	1	0	1	0	x
44	1	0	1	0	1	1	
54	1	1	0	1	0	1	x

Table C4: Truth table for successful implementation (SUC)

Row Nr.	CUST	TRACT	CENT	ENF	COH	RES	OUT	n	Raw cons.	PRI	% DCCK
13	0	0	1	1	0	0	1	2	0.852	0.746	0
29	0	1	1	1	0	0	1	3	0.826	0.655	33
59	1	1	1	0	1	0	1	1	0.791	0.447	0
62	1	1	1	1	0	1	1	2	0.774	0.522	0
14	0	0	1	1	0	1	1	1	0.760	0.603	0
32	0	1	1	1	1	1	1	1	0.715	0.357	0
16	0	0	1	1	1	1	0	1	0.714	0.278	100
15	0	0	1	1	1	0	1	3	0.709	0.370	0
51	1	1	0	0	1	0	1	1	0.707	0.417	0
64	1	1	1	1	1	1	0	2	0.704	0.311	50
3	0	0	0	0	1	0	1	1	0.702	0.386	0
61	1	1	1	1	0	0	0	6	0.674	0.409	67
48	1	0	1	1	1	1	0	5	0.666	0.344	60
39	1	0	0	1	1	0	0	4	0.658	0.472	0
52	1	1	0	0	1	1	0	3	0.657	0.478	33
18	0	1	0	0	0	1	0	2	0.655	0.370	50
45	1	0	1	1	0	0	0	2	0.642	0.320	100
11	0	0	1	0	1	0	0	1	0.634	0.236	100
27	0	1	1	0	1	0	0	4	0.630	0.270	75
23	0	1	0	1	1	0	0	1	0.600	0.317	0
40	1	0	0	1	1	1	0	3	0.596	0.365	33
46	1	0	1	1	0	1	0	3	0.595	0.325	67
19	0	1	0	0	1	0	0	1	0.590	0.069	100
55	1	1	0	1	1	0	0	3	0.586	0.344	33
17	0	1	0	0	0	0	0	1	0.558	0.248	100
31	0	1	1	1	1	0	0	5	0.545	0.168	100
56	1	1	0	1	1	1	0	2	0.531	0.186	100
2	0	0	0	0	0	1	0	1	0.530	0.276	100
63	1	1	1	1	1	0	0	9	0.529	0.201	89
47	1	0	1	1	1	0	0	6	0.529	0.146	100
33	1	0	0	0	0	0	0	2	0.509	0.373	5
50	1	1	0	0	0	1	0	4	0.502	0.119	100
49	1	1	0	0	0	0	0	4	0.501	0.275	75
34	1	0	0	0	0	1	0	5	0.332	0.101	100

Limited diversity: 34 out of 64 truth table rows are observed empirically (53.1%). Limited diversity is due to clustered, not arithmetic remainders.

Raw consistency threshold 0.7, rows 16 and 64 are excluded. Directional expectations see Table 1.

Untenable assumptions:

Parsimonious solution for SUC: tract*CENT*enf + cust*tract*COH*RES + TRACT*cent*ENF*RES + CUST*tract*CENT*res + TRACT*CENT*ENF*COH*res + cust*TRACT*cent*enf*res + CUST*cent*coh*RES

Conservative solution:

M1 : cust*CENT*ENF*coh*res + cust*tract*CENT*ENF*coh + cust*tract*CENT*ENF*res +

CUST*TRACT*enf*COH*res + cust*tract*cent*enf*COH*res + cust*TRACT*CENT*ENF*COH*RES + CUST*TRACT*CENT*ENF*coh*RES → suc (solution consistency 0.786, solution coverage 0.576).

Parsimonious solutions (without excluding untenable assumptions):

Excessive model ambiguity (36 different models, see R replication code).

Enhanced parsimonious solutions (excluding untenable assumptions):

The present data display tied logically redundant prime implicants, and hence, a modest degree of model ambiguity. All solutions are reported below (Baumgartner and Thiem 2015; Schneider and Wagemann, 2012: 108ff). The two enhanced parsimonious models differ only in path 6 regarding the role of cust versus enf*COH respectively. M1 is preferable since it has both higher consistency and coverage.

M1: cust*CENT*ENF*coh + cust*tract*ENF*res + TRACT*CENT*coh*RES + cust*TRACT*CENT*RES + CUST*TRACT*enf*COH*res + (cust*tract*cent*res) → suc (solution consistency 0.773, solution coverage 0.625).

M2: cust*CENT*ENF*coh + cust*tract*ENF*res + TRACT*CENT*coh*RES + cust*TRACT*CENT*RES + CUST*TRACT*enf*COH*res + (tract*cent*enf*COH*res) → suc (solution consistency 0.772, solution coverage 0.624).

Intermediate solutions:

M1: cust*CENT*ENF*coh + cust*tract*ENF*res + TRACT*CENT*coh*RES + cust*tract*cent*res + cust*TRACT*CENT*RES + CUST*TRACT*enf*COH*res → suc (solution consistency 0.773, solution coverage 0.625).

M2: cust*CENT*ENF*coh + cust*tract*ENF*res + TRACT*CENT*coh*RES + cust*TRACT*CENT*RES + CUST*TRACT*enf*COH*res + (cust*tract*cent*COH*res) → suc (solution consistency 0.772, solution coverage 0.624).

M3: cust*CENT*ENF*coh + cust*tract*ENF*res + TRACT*CENT*coh*RES + cust*TRACT*CENT*RES + CUST*TRACT*enf*COH*res + (tract*cent*enf*COH*res) → suc (solution consistency 0.772, solution coverage 0.624).

The three intermediate models differ only in one path regarding the role of cust, cust*COH and enf*COH, respectively as INUS conditions in combination with tract*cent*res. M1 is preferable since it has better parameters of fit and the path concerned is more parsimonious than in the other models.

Table C5: Simplifying assumptions and easy counterfactual (suc)

Row	CUST	TRACT	CENT	ENF	COH	RES	Easy counterfactual (used for intermediate solution)
1	0	0	0	0	0	0	X
5	0	0	0	1	0	0	X
7	0	0	0	1	1	0	X
26	0	1	1	0	0	1	X
28	0	1	1	0	1	1	X
30	0	1	1	1	0	1	X
58	1	1	1	0	0	1	X

Simplifying assumptions for enhanced parsimonious solution (without simplifying untenable assumptions).