

Comparative Political Studies

Political Independence, Accountability, and the Quality of Regulatory Decision-Making

Journal:	<i>Comparative Political Studies</i>
Manuscript ID	CPS-15-0521.R3
Manuscript Type:	Original Article
Keywords:	Delegation, Independence, Accountability , Regulation, Performance
Abstract:	Recent decades have seen a considerable increase in delegation to independent regulatory agencies, which has been justified by reference to the superior performance of these bodies relative to government departments. Yet, the hypothesis that more independent regulators do better work has hardly been tested. We examine the link using a comprehensive measure of the quality of work carried out by competition authorities in 30 OECD countries, and new data on the design of these organizations. We find that formal independence has a positive and significant effect on quality. Contrary to expectations, though, formal political accountability does not boost regulatory quality, and there is no evidence that it increases the effect of independence by reducing the risk of slacking. The quality of work is also enhanced by increased staffing, more extensive regulatory powers, and spillover effects of a more capable bureaucratic system.

SCHOLARONE™
Manuscripts

Political Independence, Accountability, and the Quality of Regulatory Decision-Making

Chris Hanretty and Christel Koop

Abstract

Recent decades have seen a considerable increase in delegation to independent regulatory agencies, which has been justified by reference to the superior performance of these bodies relative to government departments. Yet, the hypothesis that more independent regulators do better work has hardly been tested. We examine the link using a comprehensive measure of the quality of work carried out by competition authorities in 30 OECD countries, and new data on the design of these organizations. We find that formal independence has a positive and significant effect on quality. Contrary to expectations, though, formal political accountability does not boost regulatory quality, and there is no evidence that it increases the effect of independence by reducing the risk of slacking. The quality of work is also enhanced by increased staffing, more extensive regulatory powers, and spillover effects of a more capable bureaucratic system.

1 Introduction

In a process beginning in the 1970s, and accelerating through the 1990s, countries across the world delegated policy competences to independent regulatory agencies (IRAs) (e.g., [Majone, 1997](#); [Gilardi, 2005](#); [Jordana et al., 2011](#)). Such agencies operate under the authority of appointed rather than elected officials, and are insulated from electoral pressure. The creation of IRAs developed unevenly across countries and policy areas, with the United States being the policy innovator, and financial regulators and competition authorities being among the first to be established, followed later by utility and social regulators ([Jordana et al., 2011](#), 1346).

Decisions to delegate to IRAs partially resulted from policy diffusion by means of emulation or the social construction of necessary institutional choices ([Gilardi 2005](#); [Jordana et al. 2011](#)). As such, the intellectual argument for their establishment in each country was never fully

1
2
3 articulated from first principles. To the extent that agency creation *was* explicitly justified, the
4 central argument was that IRAs were able to make more efficient and effective policy decisions
5 than politicians and government departments, either because of their superior expertise and
6 skills, or because of certain biases or deficiencies present in decision-making by democratically
7 elected officials (Majone 1999; Gilardi 2002).
8
9

10
11
12 Studies of IRAs have concentrated on the factors that account for formal independence, of-
13 ten looking at the role of political and market-making functional imperatives (Majone 1997;
14 Gilardi 2002; Elgie and McMenamin 2005). From a normative perspective, though, this focus
15 on initial delegation seems to miss the point. Delegation to IRAs weakens the link between
16 the exercise of state power and voters, and leads to an increase in the power of unelected
17 technocrats. The standard justification of the existence of IRAs rests on their “superior perfor-
18 mance [...] relative to the result that would be likely if elected politicians were to perform the
19 functions themselves” (Thatcher and Stone Sweet, 2002, 18). Yet, in an era in which few if any
20 maintain the Saint-Simonian belief that technocratic decision-making is inherently better than
21 democratic decision-making, the reduction of democratic legitimacy requires some evidence
22 that IRAs do in fact exercise authority in such a superior fashion. That is, we need evidence
23 on whether or not granting (more) independence to IRAs results in better quality work.
24
25
26
27
28
29
30
31
32
33

34 Such evidence is hard to come by. Quality is an inherently evaluative and synoptic concept.
35 As such, scholars who are hesitant to use evaluative criteria because of the potential of serious
36 disagreement will seek to avoid investigating quality per se. At most, they will look at discrete
37 quantifiable improvements related to stated goals within a sector. Indeed, previous research
38 has assessed how IRAs promote particular desirable outcomes, such as higher network pen-
39 etration and lower interconnect rates in telecommunications (Wallsten, 2001; Gutiérrez, 2003;
40 Edwards and Waverman, 2006), more financial stability in banking (Jordana and Rosas, 2014),
41 and greater financial leverage in a range of sectors (Bortolotti et al., 2011).
42
43
44
45
46
47
48

49 Yet, the use of outcome measures has important drawbacks. Not only are quantifiable mea-
50 sures associated with gaming and neglect of other objectives (Hood and Bevan, 2006), but they
51 are also affected by much more than the regulation involved, which makes identifying the
52 specific contribution of independence challenging. In many instances, IRA creation diffused
53 rapidly, and, in the case of network industries, was concomitant with sectoral liberalization,
54 which makes it extremely difficult to disentangle the effect of IRAs from that of other changes
55 occurring at the same time. In statistical terms, the few post-liberalization data points before
56
57
58
59
60

1
2
3 the IRA is created exert considerable leverage. In practical terms, it is difficult to imagine
4 many liberalized European markets existing without an independent sectoral regulator.
5
6

7
8 In this study, we seek to provide evidence on the link between political independence and
9 regulatory quality by focusing on one particular area of regulation: competition policy, or “the
10 set of policies and laws which ensure that competition in the marketplace is not restricted in
11 a way that is detrimental to society” (Motta, 2004, 30). Competition policy is inherent to the
12 mixed economy, aiming primarily at market correction rather than creation. This allows us
13 to minimize the risk that our findings depend critically on the way in which regulation was
14 used to create new markets after liberalization.
15
16
17
18
19

20
21 We seek to improve on previous literature in three respects. First, in line with a growing body
22 of literature in economics and political science, we view independence as a matter of degree
23 rather than as a quality which is either present or absent (see, e.g., Cukierman et al., 1992; Gi-
24 lardi, 2002; Hanretty and Koop, 2012; Selin, 2015). We accordingly reframe arguments linking
25 political independence and regulatory quality so that they make sense when independence is
26 conceived in this way.
27
28
29
30
31

32
33 Second, besides independence, we take into consideration accountability, the second most dis-
34 cussed institutional feature of regulators. Accountability provisions are relevant for all policy
35 actors, but they are particularly important for IRAs. As they combine rule-making, enforce-
36 ment and adjudication, IRAs have a special duty to give account for their actions (Majone
37 1999; cf. Scott 2000; Maggetti et al. 2013). Accountability also matters for performance. It may
38 not only reduce the likelihood of misuse and abuse of political power, but may also enhance
39 the accuracy of information gathering and analysis (Patil et al., 2014). This is why several
40 authors have emphasized the need for regulators to be accountable as well as independent
41 (e.g., Majone, 1999; Quintyn and Taylor, 2007; Busuioc, 2009). We therefore assess the impact
42 of independence as well as accountability, distinguishing the two concepts conceptually and
43 empirically, and analyzing new data on the formal political independence and accountability
44 of competition authorities.
45
46
47
48
49
50
51
52
53

54
55 Third, we use a comprehensive proxy measure of regulatory quality – in our case, the quality
56 of competition policy enforcement – which is produced by an external company (the Global
57 Competition Review or GCR) based on expert opinion and additional data. The measure
58 centres on the enforcement *process*, capturing the quality of the work done by competition
59 authorities, including their economic and legal analyses and their intermediate output. Con-
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

sequently, we are able to focus on the contribution of the authorities themselves, rather than on certain market outcomes which are affected by many other factors. The data on regulatory quality allow us to assess the impact on quality of the formal independence and accountability of competition authorities in 30 member states of the Organization for Economic Co-operation and Development (OECD), for the period 2005-2014. The results of our ordered probit regression of regulatory quality show that independence indeed improves quality. However, the effect of accountability is not in the hypothesized direction, raising questions of institutional design and the role of accountability more generally.

2 Independence and regulatory quality

Independence is the most commonly discussed characteristic of IRAs. The concept serves two purposes in the literature. First, it demarcates the scope of the genus: a regulator must be independent in some minimal sense to be considered an IRA. For Thatcher, for instance, the minimum requirements for categorization as an IRA comprise: “(1) that the agency has its own powers and responsibilities given under public law; (2) that it is organizationally separated from ministries; and (3) that it be neither directly elected nor managed by elected officials” (2005, 352; cf. Majone 1999, 2; Jordana et al. 2011, 1351). Second, the concept of independence allows us to differentiate *within* the category of IRAs. That is, independence is a matter of degree, with some regulators being more independent than others.

In this study, we focus on political independence, or the independence of regulators from elected politicians and members of government. In concentrating on political independence, we do not mean to imply that independence from other actors – and from the regulated sector in particular – is unimportant; merely that it is independence from politics that is implied by the literature, justifying the creation of IRAs. By the political independence of an agency, we mean “the degree to which the agency takes day-to-day decisions without the interference of politicians in terms of the offering of inducements or threats and/or the consideration of political preferences” (Hanretty and Koop, 2013, 196).

Political independence defined this way is a property of agencies’ behavior, or what actually happens. It is common in the literature to make a distinction between formal and actual independence, or between what is true *de jure* and *de facto*. This has been the case since the early literature on *de jure* and *de facto* central bank independence (Cukierman et al., 1992). We concentrate on formal, or *de jure* independence, which we conceptualize as the degree to

1
2
3 which there are statutory provisions that decrease the possibilities for politicians to influence
4 agency decisions before they are made. We focus on formal independence because this is
5 the feature of IRAs that can most easily be 'engineered', and because there is considerable
6 evidence linking grants of formal independence to higher degrees of actual independence,
7 at least in established democracies (e.g., [Hanretty and Koop 2013](#); [Enns-Jedenastik 2015](#)).
8 Nonetheless, our findings do not automatically translate to de facto independence.
9

10
11
12 How then should we expect formal independence and regulatory quality to be related? Al-
13 though early studies of independent regulatory commissions in the United States already
14 associated independence with higher levels of professionalism, consistency and policy con-
15 tinuity (see [Bernstein, 1955](#), Ch. 5), most theorization of the link has been done in recent
16 decades. Two main causal arguments have been put forward: an expertise-based and a credi-
17 ble commitment-based argument.
18
19

20
21
22 The argument from expertise maintains that, compared to politicians and government depart-
23 ments, independent agencies have better access to (or can better process) information, and that
24 better information leads to better work. Though bureaucrats generally have better information
25 than politicians, this does not explain why bureaucrats in an independent agency should have
26 better information than bureaucrats in a government department and drawn from a career
27 civil service. To explain this, we have to examine the act of delegation.
28
29

30
31
32 [Bawn \(1995\)](#) explains that in setting the level of agency independence, politicians also make
33 decisions about the administrative procedures agencies may employ, and these procedures
34 may limit the information available to the agency, and the ability to assess policy consequences
35 and make decisions based on expertise. Although the examples Bawn gives are drawn from
36 the study of executive agencies in the United States, initial decisions about grants of inde-
37 pendence have also affected the ability of IRAs in Europe to draw on expertise. Limitations
38 may come through obligations to consult with the more generalist and political leadership of
39 government departments, but also through restrictions on the use of outside consultancies or
40 on the hiring of short-term experts.
41
42

43
44
45 This claim has been challenged. [Gailmard and Patty \(2007\)](#) argue that the link between in-
46 dependence and expertise is conditional on bureaucratic selection and retention: only policy-
47 motivated "zealots" protected by long tenure will invest in expertise, whereas other agency
48 staff ("slackers") will not because of a lack of incentives. That is, bureaucratic expertise is
49 costly to develop and relationship-specific, while the specific relationship may not continue.
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Though these findings have some relevance for IRAs, we expect the effect of independence to be stronger for these bodies as the relevant (regulatory) expertise is less relationship-specific than is assumed by Gailmard and Patty.¹ Thus, if expertise is an ingredient of regulatory quality, the literature should lead us to expect a positive effect of independence on quality.

A somewhat more recent strand of literature focuses on the role of credible commitment in regulation (e.g., Majone 1996; Levy and Spiller 1996; Gilardi 2002). The origins of the link between independence, credible commitment and better policy-making lie in the field of monetary policy, where it is common to argue that politicians use monetary policy to boost the economy (i.e., increase employment and growth); that politicians' actions in this regard are rationally expected and anticipated by price- and wage-setters; and that as a result, discretionary monetary policy leads only to above-target inflation rather than to the desired gains in employment and output (Kyland and Prescott, 1977). To avoid inflationary bias, and to deal with the time inconsistency in monetary policy, politicians may wish to tie their hands, and delegate to an independent and more 'conservative' central bank (Rogoff, 1985).

Regulation scholars have argued that time inconsistency is important in their field, too. Politicians wish to commit to long-term regulatory policies, but face incentives to deviate and pursue politically attractive short-term policy options (Majone, 1996). For instance, while price caps are regulatory instruments to correct market failure in monopolistic or oligopolistic sectors, they may also be used by politicians for electoral purposes. Yet, anticipating such use of caps, potential investors may stay or move away from these sectors, leading to markets being hampered rather than corrected (cf. Levy and Spiller 1996). By delegating to IRAs headed by officials with different time horizons and incentives, politicians can tie their hands and commit credibly to long-term regulatory objectives and better decisions.

Summarizing the arguments, agency independence may sever the link between the political preferences and policy decisions – a point we come back to in the next section – but it enhances regulatory quality, whether we follow the logic of expertise or the logic of commitment. Hence, we hypothesize that:

H1. The higher the degree of formal political independence, the higher the quality of regulatory decision-making.

¹ That is, looking at the 'revolving door' literature, regulatory expertise seems rather transferable to the private (regulated) sector (Moe 2012, 1171.)

3 Accountability and regulatory quality

Political accountability plays a crucial role in normative discussions of IRAs. That is, if regulators are not accountable to anyone, their legitimacy is called into question (Majone 1999; Scott 2000). Accountability does not feature prominently in empirical studies though. The main reason is that many scholars equate it with control, which is itself considered to be the inverse of independence. Hence, mechanisms designed to ensure accountability by definition compromise independence, and assessing both independence and accountability becomes nonsensical.² Yet, a closer look at the two concepts may lead us to treat them as separate dimensions.

Scholars working on accountability typically define the concept in terms of answerability, or the ability to provide information on, and explanation of, one's conduct.³ For Philp, for instance, "A is accountable with respect to M when some individual, body or institution, Y, can require A to inform and explain/justify his or her conduct with respect to M" (2009, 32; cf. Scott 2000, 40). Analogous to our treatment of independence, we concentrate on formal accountability to politicians, acknowledging all the while that the actual use of accountability may not correspond to what provisions prescribe, and that other forms of accountability (horizontal and downwards) may also matter. We say that an IRA is formally accountable to politicians to the extent that politicians can require the agency to provide information on, and explanation of, its conduct on the basis of statutory provisions rather than non-legal forms of compulsion.

Defined in this way, the independence of IRAs is compatible with some degree of accountability. That is, agencies may make their day-to-day decisions independently from politicians and political preferences, but may still be required to provide information on, and explanation of, these decisions ex post facto. Indeed, recent empirical research has shown that agencies can be both independent and accountable – they can "have their cake and eat it, too" (Maggetti

² Something like this belief is certainly implicit in a number of coding schemes used to operationalize political independence, with reporting requirements being treated as inimical to independence (e.g., Gilardi 2002; Edwards and Waverman 2006).

³ This understanding is also reflected in dictionary definitions. For instance, the Oxford English Dictionary defines accountable as "required or expected to justify actions or decisions." The possibility of sanctions is a second component that is frequently included, but there is no agreement on this element. As Philp sets out, sanctions may enable accountability to better achieve certain outcomes, but this may be a potential additional feature rather than an attribute (2009, 35-36).

1
2
3 *et al., 2013*).⁴ Thus, accountability and independence need to be operationalized carefully and
4 separately, and their effects disentangled.
5
6

7
8 Systematic empirical research on the impact of accountability on policy-making is rare. As
9 *Dubnick and Frederickson (2011)* put it, the literature has largely failed to go beyond ‘promises’
10 of what accountability can achieve. Yet, there are two literatures that explicitly link the two
11 concepts: the constitutionalist and the social-psychological literature.
12
13

14
15 For constitutionalist scholars, first of all, the role of accountability – and checks and balances
16 more generally – is one of preventing and detecting abuse and misuse of political power.
17 Constitutionalism is grounded in a considerable distrust of power-holders. As *Loewenstein*
18 put it: “A stigma is attached to power, and only the saints among the power holders – rarely
19 found – are able to resist the temptation of abusing it” (1957, 8). The solution to the problem
20 is believed to lie in institutional design, which would allow us to constrain and restrain the
21 exercise of political power.
22
23
24
25
26
27

28
29 These ideas have found expression in the Federalist Papers and the US Constitution, and they
30 are at the heart of constitutionalist principles such as limited government, checks and bal-
31 ances, separation of powers and judicial review. They are also reflected in the literature on
32 the role of accountability arrangements. That is, requirements to provide information and
33 justification are considered to help prevent and detect arbitrary exercise and abuse of public
34 authority, including the misuse of public funds, the pursuance of particular rather than gen-
35 eral interests, the unequal or unfair treatment of citizens (or companies), the neglect of rights
36 and freedoms, and corruption and patronage (e.g., *Bovens, 2007*). This is not only important
37 from a normative point of view, but it may also effect organizational performance. If political
38 power-holders such as regulatory agencies may be tempted to misuse and abuse their pub-
39 lic authority, if these activities are associated with worse performance, and if accountability
40 mechanisms may provide incentives not to engage in these activities, we shall expect higher
41 degrees of accountability to be associated with higher quality of (regulatory) decision-making.
42
43
44
45
46
47
48
49
50
51

52
53 The micro-foundations of the link between accountability and performance can be found in the
54 field of social psychology and, in particular, the work by Tetlock (see *Patil et al., 2014*). What
55 is stressed in this literature is the role of accountability as a generator of implicit or explicit
56 expectations “that one may be called on to justify one’s beliefs, feelings and actions to others”
57
58

59
60 ⁴ This is also how practitioners tend to think about the concepts. Looking at the websites of prominent regu-
lators and central banks, one finds statements about their independence as well as their accountability. Also, an
expert meetings on regulation organized by the OECD in January 2005 was titled “Designing Independent and
Accountable Regulatory Authorities for High Quality Regulation”.

1
2
3 (Lerner and Tetlock, 1999, 255). These expectations play an important role in decision-making
4 as they lead actors to be less prone to the fundamental attribution error (Tetlock, 1985), and
5 to process information more carefully (Tetlock, 1983) and accurately (Mero and Motowidlo,
6 1995).
7
8
9

10
11 The link between accountability and the quality of decision-making has mainly been estab-
12 lished at the individual level. As our research focuses on organizations, we need to be careful
13 when formulating our expectations: organizational decision-making cannot be treated as a
14 simple aggregate of individual-level processes. Yet, there is at least some evidence suggesting
15 that the effect of accountability on the use of information operates at the individual as well as
16 the group level (Weldon and Gargano, 1988).
17
18
19
20
21

22
23 However, studies of accountability do not unequivocally point to a positive effect on perfor-
24 mance. Accountability is also associated with excessive costs and red tape. This is particularly
25 true for provisions that are generic in nature, including standardized and routine requirements
26 such as obligations to produce annual reports and evaluation protocols (Bovens and Schille-
27 mans, 2014). Indeed, scholars have emphasized how sensitive the design of accountability is,
28 and how important appropriate design is for accountability to have desirable effects on per-
29 formance (Quintyn and Taylor, 2007; May, 2007). Nonetheless, as accountability provisions
30 are mainly linked to better performance, we formulate the following hypothesis:
31
32
33
34
35
36

37 *H2.* The higher the degree of formal political accountability, the higher the quality of
38 regulatory decision-making.
39
40
41

42 Accountability may not only have a direct effect on the quality of regulatory decision-making,
43 but may also strengthen the effect of independence. For this argument to make sense, we need
44 to have a closer look at the differential effect of independence. As set out in Section 2, political
45 independence can increase the level of expertise in regulatory decision-making, and reduce
46 short-termism in the process. Yet, insulating regulatory agencies from politicians comes at a
47 cost: by granting higher degrees of political independence, one also enhances the possibility
48 of slack and bureaucratic drift (e.g., McCubbins et al. 1987; Bawn 1995; Gailmard and Patty
49 2007). That is, regulators may actually do something other than that which they were initially
50 directed to do by their political principals.
51
52
53
54
55
56
57

58 Partially, political independence is precisely intended to create the possibility for regulatory
59 agencies to drift. Following the logic of credible commitment, regulators shall stick to long-
60 term policy objectives even if their political principals want them to (temporarily) move away

1
2
3 from these objectives. Hence, the potential decline in policy responsiveness to politicians'
4 preferences may constitute an asset rather than a cost. However, policy unresponsiveness
5 is not the only concern; bureaucratic drift may also take a more procedural form. Political
6 independence removes some checks and balances from the policy process as it reduces the
7 opportunities for politicians to keep an eye on the way in which agencies carry out their
8 work. Reduced oversight makes it, *ceteris paribus*, easier for regulators to exercise their power
9 arbitrarily and to misuse their funds. It may result in more corner-cutting, fewer procedural
10 checks, more superficial investigations of complaints, and lower levels of due process. In sum,
11 political independence raises concerns of a constitutionalist nature.
12

13
14 Ideally, regulators are designed in such a way that independence enhances levels of exper-
15 tise and long-termism, while opportunities for misuse and abuse of power are being curbed.
16 Such a design may be created by complementing political independence with provisions for
17 accountability (Majone, 1999; Quintyn and Taylor, 2007). If the two concepts are, indeed,
18 “complementary and mutually supporting” (Majone, 1999, 14), accountability can serve as a
19 guard against abuse and misuse, as regulators are called upon to explain and justify their
20 actions – or lack thereof (cf. Quintyn and Taylor, 2007, 35). Accordingly, we hypothesize that
21 the effect of independence will depend on the level of accountability, and more specifically:
22

23
24
25
26
27
28
29
30
31
32
33
34
35 *H3.* The higher the level of formal political accountability, the stronger the effect of
36 formal political independence on the regulatory quality of agencies
37

38 39 40 41 **4 Data and operationalization**

42
43 Having hypothesized the relationship between independence, accountability and the qual-
44 ity of regulatory decision-making, we now turn to the operationalization of these and other
45 variables.
46
47
48

49 50 51 **Independence and accountability**

52
53 Our measures of formal political independence and accountability are based on the analysis
54 of the statutory provisions governing independent competition authorities in the 30 OECD
55 countries included in our analysis. As we are interested in the variation in the design of
56 *independent* regulatory agencies, we excluded from our analysis those competition authorities
57
58
59
60

1
2
3 that were part of a ministerial hierarchy in any given year.⁵ Thus, the Antitrust Division of the
4 United States Department of Justice was excluded, as were the Belgian Directorate General for
5 Competition (operating until 2013), Spain's Competition Service (operating until 2007), and
6 the non-independent Dutch Competition Authority (operating until 2005).
7
8
9

10
11 Over the period in question (2005-2014), we observed considerable within-country (and within-
12 agency) variation in design. Incorporating this variation in the analysis is important as it al-
13 lows for more controlled comparison than does the cross-country variation. We therefore dis-
14 aggregated competition authorities, and considered competition authorities which had been
15 reformed, or which have been subjected to alterations in their independence or accountability,
16 as separate bodies. Thus, although we have data for thirty countries, we have information for
17 46 regulators, or regulator-spells.⁶
18
19
20
21
22
23

24 For each authority, we recorded information regarding different items relating to their statu-
25 tory design. The selection of the formal political independence items was based on previous
26 literature, particularly the work by Gilardi (2002). For each item, we coded the agency's 're-
27 sponse', or the category that best described the statutory provisions imposed upon the agency.
28 For some items – for example, provisions which are either present or absent, and which are
29 commonly found in statutes governing competition authorities – the coding was relatively
30 straightforward. For other items, though, the coding was more difficult, either because the
31 provision found in the statute was not immediately reconcilable with the list of options, or
32 because the relevant provision was a matter of administrative law more generally. The coding
33 was carried out by one of the authors and a research assistant; differences in coding were
34 discussed by the authors and reconciled by mutual agreement.
35
36
37
38
39
40
41
42
43

44 To assess the scalability of the items, and to turn our ordinal measurements on several items
45 into a single measure, we used a latent trait model based on item response theory (IRT) (cf.
46 Hanretty and Koop, 2012; Selin, 2015). IRT models treat the observed item responses as a
47 function of an unobserved latent variable – in our case, the trait we refer to as independence.
48 They allow for the estimation of the weights of items, and the scores of item responses, on
49 the latent trait, and for the exclusion of those items for which reasonable weights and scores
50
51
52
53
54

55 ⁵ One could, in principle, compare independent and non-independent competition authorities. However, it
56 would be inappropriate to give non-independent authorities (or ministerial divisions) the value of zero on the
57 independence index as being a non-independent authority is not the same as being an independent authority
58 with the lowest value on the index. The alternative is to compare the two categories of organizations, but this
59 is problematic because there are only a few non-independent competition authorities for which we have data on
60 quality.

⁶ Only one country – the UK – had two independent competition authorities (the Competition Commission
and the Office of Fair Trading), and only in the period before 2014. We included both.

cannot be estimated. IRT models are similar to confirmatory factor analysis in the sense that both attempt to identify an underlying latent factor. Yet, while factor analysis requires the inclusion of continuous data, IRT models do not have such a requirement (Raju et al., 2002, 520). This makes IRT models more appropriate for our data structure, which includes items that are measured categorically and even dichotomously. Our aggregate measure of independence is the measures that best ‘explains’ the observed pattern of responses, given a set of assumptions.

Formally speaking,⁷ suppose we have information for J regulators on $I = 1, \dots, I$ items, each of which has K categories. These items may ‘discriminate’ between regulators to a greater or lesser extent. Use β_i to stand for the discrimination parameter, and interpret it as analogous to a factor loading in a conventional factor analysis (see ‘discrimination’ in Tables 1 and 2). In this way, an item with an associated value of β equal to 10 ‘matters more’ for independence than an item with a β value of 1. Within each item, the individual response categories may be ‘easier’ or ‘more difficult’ to achieve, depending on the regulator’s level of independence. Use α_{ik} to refer to the location of each item response category, such that response categories with higher values of alpha require ‘more’ of the latent trait (see ‘location’ in Tables 1 and 2). If we use θ_j to stand for the level of the latent trait, then we can model the manifest responses as follows:

$$\text{logit}(\gamma_{ik}) = \beta_i(\theta_j - \alpha_{ik}^*),$$

where γ_{ik} is the cumulative probability of a response in the k th category or higher to the i th item. Thus, although the left-hand side of the equation features the manifest responses of each regulator, the interest lies in the parameter on the right hand side, namely θ_j , the measure of the formal independence of each competition authority. Whether or not a regulator manages to attain a particular category within an item depends on the gap between its latent level of independence (θ) and the difficulty of the item category (α), as well as the discrimination parameter (high independence regulators might not demonstrate a high-independence response if the discrimination of that item is low).

Our final measure of independence includes all items originally used by Gilardi (2002), with three exceptions. First, we excluded items that are conceptually linked to the accountability and powers of agencies rather than to independence. These are also items which previous

⁷ This section describes Samejima’s graded response model (Samejima, 1970), though note that we reverse the symbols used to describe the location and discrimination parameters.

research on IRAs demonstrated to be unrelated to the latent trait (see Hanretty and Koop, 2012). Second, we excluded items which turned out to be unrelated to the trait of formal independence, thus excluding provisions on the identity of appointing actors. Our findings in this respect are largely in line with those in previous work on a larger sample of independent regulatory agencies (Hanretty and Koop, 2012). Third, we excluded items which lacked variation in our specific sample of competition authorities. We were left with eighteen items in total. These items, the response categories and the latent trait data are shown in Table 1.⁸

Table 1: *Independence items*

Item	Categories	Location	Discrimination
What type of position does the agency head have?	Permanent position		1.1
	Fixed-term position	- 3.9	1.1
What is the tenure of the head of the agency?	4 years or fewer		0.2
	5 years	- 1.0	0.2
	6 years	0.1	0.2
	More than 6 years	1.2	0.2
Do the founding documents give provisions relating to the dismissal of the head?	No, there are no such provisions		225.5
	Yes, there are such provisions	-658.6	225.5
Under what conditions may the head be dismissed?	Head may be dismissed for any reason		0.9
	Head may be dismissed only for non-political reasons (illness, incapacity)	- 2.4	0.9
	Head may not be dismissed	- 1.0	0.9
Do the founding documents give provisions relating to the incompatibility of the head?	No, there are no such provisions		0.9
	Yes, there are such provisions	- 3.3	0.9
May the appointment of the head be renewed?	Yes, the appointment may be renewed indefinitely		
	Yes, the appointment may be renewed once only		
	No, the appointment may not be renewed		
Is there a formal statement of the independence of the agency head?	No		0.1
	Yes	0.4	0.1
What is the tenure of board members?	4 years or fewer		1.3
	5 years	0.3	1.3
	6 years	1.4	1.3
	More than 6 years	2.5	1.3
Under what conditions may board members be dismissed?	Members may be dismissed for any reason		0.6

⁸ The list includes an item on politicians being able to overturn the decisions of the competition authority, which some jurisdictions allow for. Importantly, it is the judiciary that is the main branch when it comes to vetoing decisions. However, not only does the latter capture independence from the judiciary rather than from politics, it is also a feature on which no variation exists. That is, in all countries, affected parties can turn to the judiciary to appeal the competition authority's decisions.

Table 1: (continued)

Item	Categories	Location	Discrimination
	Members may be dismissed only for non-political reasons (illness, incapacity)	- 1.8	0.6
	Members may not be dismissed		0.6
Do the founding documents give provisions relating to the incompatibility of board members?	No, there are no such provisions		
	Yes, there are such provisions		
May the appointment of board members be renewed?	Yes, the appointment may be renewed indefinitely		
	Yes, the appointment may be renewed once only		
	No, the appointment may not be renewed		
Is there a formal statement of the independence of board members?	No		0.9
	Yes	1.5	0.9
Is there a formal statement of the agency's independence?	No		0.3
	Yes	- 1.4	0.3
May the agency's decisions be overturned by a politician?	Yes		
	Only in exceptional cases		
	No		
What is the source of the agency's budget?	Only fees/levies from the industry		0.7
	Both fees/levies and government funding	2.0	0.7
	Only or primarily government funding	3.3	0.7
How is the budget controlled?	By the agency itself (this is pretty much always the case; also if (2) or (3))		0.0
	By an audit office, accountant and/or court	- 1.0	0.0
	By the government and/or parliament/parliamentary committee (and typically also still (2) in that case)		0.0
Which body decides on the agency's internal organisation	By the agency alone		0.0
	By the agency and the government	- 0.8	0.0
	By the government alone	0.6	0.0
Which body decides on the agency's personnel policy?	By the agency alone		
	By the agency and the government		
	By the government alone		

We followed the same procedure to create an aggregate measure of formal political accountability. Again, the selection of these items was based on previous literature. Specifically, we followed Koop (2011), except that we excluded items which lacked variation in our sample

of competition authorities. The remaining items relating to accountability, and the different responses to these items, are shown in Table 2. We were left with eleven items in total.

Table 2: *Accountability items*

Item	Categories	Location	Discrimination
Is the agency required to send information upon request to government?	No		0.8
	Yes	-1.3	0.8
Is the agency required to send information upon request to parliament?	No		0.6
	Yes	0.5	0.6
Is the agency required to submit an annual plan to government?	No		0.7
	Yes	1.0	0.7
	Yes, and it also needs approval	2.4	0.7
Is the agency required to submit an annual or multi-annual plan to parliament?	No		0.6
	Yes	0.9	0.6
	Yes, and it also needs approval	2.3	0.6
Is the agency required to submit an annual itemised budget to government?	No		0.4
	Yes	0.1	0.4
	Yes, and it also needs approval	1.5	0.4
Is the agency required to submit an annual itemised budget to parliament?	No		0.7
	Yes	1.0	0.7
	Yes, and it also needs approval	2.4	0.7
Is the agency required to submit an annual activity report to government?	No		0.3
	Yes	-1.2	0.3
	Yes, and it also needs approval	0.2	0.3
Is the agency required to submit an annual activity report to parliament?	No		0.2
	Yes	-0.8	0.2
	Yes, and it also needs approval	0.6	0.2
Is the agency required to submit an annual financial report to government?	No		0.7
	Yes	-0.6	0.7
	Yes, and it also needs approval	0.8	0.7
Is the agency required to submit an annual financial report to parliament?	No		0.7
	Yes	0.0	0.7
	Yes, and it also needs approval	1.4	0.7
Can parliament hear the head of the agency?	No		0.5
	Yes	-1.0	0.5

The aggregate measures are shown in Table 3, which lists all authorities in alphabetical order.


The table also shows the limited association between the formal independence of authorities and their formal accountability towards politicians: the Pearson correlation coefficient between the two variables measured the agency level is only $r = -0.16$, which is not significant at standard levels (d.f. = 29, $p = 0.39$).

Table 3: *Independence, accountability and ratings of agencies over time*

Country (Period)	Rating	Independence	Accountability
AUS (2004–2010) <i>Australian Competition and Consumer Commission</i>	4 → 4	1.41	0.91
AUS (2011–2014) <i>Australian Competition and Consumer Commission</i>	4 → 4	1.41	0.91
AUT (2004–2014) <i>Federal Competition Authority</i>	3 → 3	-1.38	-0.79
BEL (2004–2005) <i>Competition Council</i>	2.5 → 3	0.81	-1.29
BEL (2006–2012) <i>Competition Council</i>	2.5 → 2.5	0.81	-1.29
BEL (2013–2014) <i>Belgian Competition Authority</i>	2.5 → 2.5	0.91	-0.33
CAN (2004–2014) <i>Competition Bureau</i>	4 → 3.5	-1.43	-0.93
CHE (2004–2014) <i>Competition Commission</i>	3 → 3	0.7	-1.67
CHL (2008–2014) <i>National Economic Prosecutor's Office</i>	2 → 3	-1.51	-0.21
CZE (2006–2014) <i>Office for the Protection of Competition</i>	3 → 2.5	-1.39	0.65
DEU (2004–2014) <i>Federal Cartel Office</i>	4 → 5	-1.44	-0.99
DNK (2004–2014) <i>Competition and Consumer Authority</i>	4 → 2.5	-1.43	-0.52
ESP (2008–2012)	3.5 → 4	0.89	1.38

1				
2				
3				
4	<i>National Competition Commission</i>			
5	ESP (2013–2014)	4→4	1.56	1.38
6				
7	<i>National Markets and Competition Commission</i>			
8				
9	FIN (2004–2012)	3.5→3	-1.45	0.55
10				
11	<i>Competition Authority</i>			
12				
13	FIN (2013–2014)	3→3	-1.38	0.55
14				
15	<i>Competition and Consumer Authority</i>			
16				
17	FRA (2004–2008)	4→4	-0.06	-1.29
18				
19	<i>Competition Council</i>			
20				
21	FRA (2009–2014)	4→5	-0.09	-1.29
22				
23	<i>Competition Authority</i>			
24				
25	GBR (2004–2013)	4.5→4.5	-0.25	-0.23
26				
27	<i>Competition Commission</i>			
28				
29	GBR (2004–2013)	4→4	-0.39	0.03
30				
31	<i>Office of Fair Trading</i>			
32				
33	GBR (2014–2014)	4	-0.39	0.03
34				
35	<i>Competition and Markets Authority</i>			
36				
37	GRC (2004–2011)	2→2.5	0.08	0.12
38				
39	<i>Competition Commission</i>			
40				
41	GRC (2012–2014)	3→3.5	1.25	0.22
42				
43	<i>Competition Commission</i>			
44				
45	HUN (2006–2014)	3→3	0.52	-0.61
46				
47	<i>Competition Authority</i>			
48				
49	IRL (2004–2013)	4→3	-0.29	0.53
50				
51	<i>Competition Authority</i>			
52				
53	IRL (2014–2014)	2.5	-0.29	0.66
54				
55	<i>Competition and Consumer Protection Commission</i>			
56				
57	ISR (2004–2011)	3→3	-1.39	-1.8
58				
59	<i>Antitrust Authority</i>			
60				
	ISR (2012–2014)	3→3	-1.39	-1.8
	<i>Antitrust Authority</i>			
	ITA (2004–2014)	4→3.5	1.66	-0.23

1
2
3
4 *Authority for Competition and the Market*

5 JPN (2004–2014)  1.31 -1.4

7 *Fair Trade Commission*

9 KOR (2004–2014)  0.02 -1.8

11 *Fair Trade Commission*

13 MEX (2004–2012)  1.47 0.83

15 *Federal Competition Commission*

17 MEX (2013–2014)  2.17 1.23


19 *Federal Economic Competition Commission*

21 NLD (2005–2012)  -0.1 -0.8


23 *Netherlands Competition Authority*

25 NLD (2013–2014)  0.38 -0.21

27 *Authority for Consumers and Markets*

29 NOR (2004–2014)  -1.44 -0.55

31 *Competition Authority*

33 NZL (2004–2014)  -0.09 1

35 *Commerce Commission*

37 POL (2004–2006)  -1.56 0.1

39 *Office of Competition and Consumer Protection*

41 POL (2007–2014)  -3.52 0.1

43 *Office of Competition and Consumer Protection*

45 PRT (2004–2013)  0.91 1.19

47 *Competition Authority*

49 PRT (2014–2014)  1.73 1.39

51 *Competition Authority*

53 SVK (2006–2011)  -1.39 1.2

55 *Antimonopoly Office of the Slovak Republic*

57 SWE (2004–2008)  -1.42 -0.55

59 *Competition Authority*

61 SWE (2009–2014)  -1.42 -0.55

63 *Competition Authority*

65 TUR (2011–2014)  1.48 -0.2

Competition Authority

USA (2004–2014)

4.5  5

1.52

0.56

*Federal Trade Commission***Regulatory quality**

There is a broad literature on ‘regulatory quality’. This is due in part to the multidimensionality of the concept. Indeed, ‘high-quality’ regulatory work refers to work that is efficient, proportionate, legitimate, consistent, not unduly prescriptive, and enforceable. Yet, even if we agree on the characteristics, the measurement of quality, or of its constituent characteristics, remains difficult. Researchers have generally accepted overall measures of regulatory quality at the country level, as is seen from the widespread use of World Bank Governance indicators (Kaufmann et al., 2010). It is both possible and desirable to seek out measures that are finer-grained both in terms of their level of application (organizational rather than country level), and in terms of the facets of quality they examine. Though the GCR’s comprehensive proxy measure of the quality of competition policy enforcement has, as we will point out, its weaknesses, it comes closest to what we consider desirable.

Three approaches to quality in regulation can be identified in the literature: those focusing on *process*, *activity* and *real-world outcomes* (Radaelli, 2004). Process-based approaches conceive of regulatory quality as a characteristic of the work of regulators such that the procedures, investigations, analyses, and intermediate outputs are of good quality, judged according to standards internal to the kind of work. Thus, high-quality work may refer to well-targeted investigations, clear and well-informed legal analyses, and state-of-the-art econometric analyses. Activity-based approaches, on the other hand, understand quality as a characteristic of regulatory work such that regulators doing more work (e.g. more cartel investigations, more sectoral inquiries) are of better quality. Finally, outcome-based approaches take quality to mean regulatory work that promotes those outcomes that regulators are meant to promote.

Outcome-based approaches have so far be most common. Yet, as set out in Section 1, outcomes are affected by many other variables for which it is hard to control. Activity-based approaches also have their problems: in the case of competition authorities, high levels of activity may either imply that the agency is doing a good job – i.e., it manages to detect non-competitive

1
2
3 behavior – or that it is failing to do its job well – i.e., it has not managed to create competitive
4 markets. Moreover, to the extent that certain activities are included in targets, levels of activity
5 may be affected by gaming, with other, non-measured activities being neglected. For those
6 reasons, we rely on a measure which is primarily process-based. We concentrate on high
7 quality processes, as opposed to high levels of activity or high-quality outcomes (though it
8 would be surprising if high quality processes persistently failed to deliver good outcomes).
9

10
11
12
13
14
15
16 Whereas our measures of independence and accountability are based on collected data, our
17 measure of the quality is based on what might be termed ‘found data’. The GCR has, since
18 2000, published an annual report called *Rating Enforcement*. This rates the work of competi-
19 tion agencies from across the world in the past year. The first edition rated 16 agencies in
20 13 countries, with the United States being the only non-European country. The 2015 edition
21 which we use rates 36 agencies from 34 countries, including IRAs from Asia and Latin Amer-
22 ica ([Global Competition Review, 2015](#)). We use these ratings as the basis for our analysis,
23 excluding countries from non-OECD countries.⁹
24
25
26
27
28
29
30

31
32 The ratings published by the GCR are produced by combining the results of an expert survey
33 with quantitative data and qualitative judgements made by the GCR editorial team. The
34 methods used to produce these ratings have changed over time. Since 2005, ratings have been
35 ascribed to individual agencies rather than to competition regimes, which had happened in
36 the years before. We therefore use the ratings for the years 2005 to 2014. Experts are invited
37 to give an aggregate star rating to agencies with which they are familiar (thus permitting –
38 and indeed encouraging – experts to judge agencies from different countries), and are then
39 invited to give open-ended responses to questions on the regulator(s).
40
41
42
43
44
45
46

47
48 GCR describes the expert survey as a “user’s survey”, defining a user of a competition au-
49 thority as anyone with “cause to liaise with a competition agency on a proceeding whether
50 as a private practitioner, an in-house counsel, business executive, consumer representative, or
51 as an economist”, though it notes that the majority of respondents tend to be lawyers and
52 economists ([Global Competition Review, 2005, 2](#)). The list of respondents was constructed us-
53 ing trade publications, including but not limited to *The International Who’s Who of Competition*
54 *Lawyers*, *The International Who’s Who of Competition Economists*, *40 under 40*, *Women in Antitrust*,
55
56
57
58
59
60

⁹ The reason is that the factors promoting quality in established democracies may be rather different from those in developing democracies.

1
2
3 and *The Who's Who of Public Affairs*. Subscribers to the service were also given the possibility
4 to fill out the survey. For identified experts, the response rate was approximately one in seven
5
6
7 ([Global Competition Review, 2006, 2](#)). The eventual ratings depend on the results of this ex-
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

and *The Who's Who of Public Affairs*. Subscribers to the service were also given the possibility to fill out the survey. For identified experts, the response rate was approximately one in seven ([Global Competition Review, 2006, 2](#)). The eventual ratings depend on the results of this expert survey, and on the opinion of GCR staff, who conduct in-depth interviews and analyses of statistical data on budgets, case loads, staffing, and the reporting in the *Global Competition Review* itself. These opinions are included to introduce 'flexibility', and to make ratings more comparable across countries ([Global Competition Review, 2011, 11](#)).

Though we do not know exactly what it is about the enforcement of competition policy that discriminates between the different ratings, the GCR does provide insight into this question, and this information leads us to conclude that the notion of quality captured by the GCR is related to a process-based understanding of quality, rather than a real-world outcomes or activity-based approach. First, the ratings aim to capture the performance of the agencies and specifically disavows consideration of outcomes in terms of the success of competition policy or the performance of economy ([Global Competition Review, 2008](#)). Second, the ratings primarily incorporate assessments of the overall quality of the process of enforcement rather than the output. Consequently, competition authorities that only take decisions on cases that are brought before them by another authority — tribunal-like bodies such as the Canadian Competition Tribunal and the Finnish Market Court — are excluded. The latter are considered 'passive' authorities, while the GCR is interested in the 'active' part of enforcement — the monitoring and the investigations. Although the GCR emphasizes that it collects "a great deal of data on agency activity", including data on merger, cartel and abuse of dominance activity and policy and advocacy work ([Global Competition Review, 2005, 2](#)), these statistics are only ever read in conjunction with the questions in the surveys which ask users to evaluate, for each agency they are familiar with, its areas of excellence, the area that could improve, the margin of improvement over the past five years, markets which the agency does and does not understand, the suitability of the agency head to the agency's needs, and the agency's morale. While we can never be sure about the motivation of users when they rated competition authorities – which is a more general problem of (inter-)subjective measures–, the fact that they were asked to reflect on these specific issues makes it at least more likely that they rated quality rather than other features which they were (dis)satisfied with.

The GCR also sets out what type of factors contribute to higher ratings ([Global Competition](#)

1
2
3 Review, 2005, 2007). Features that play such a role are (a) vigorously pursuing in all areas the
4 agency is responsible for, including politically sensitive areas such as domestic cartels, abuse
5 of dominance – particularly by former state-owned companies – and government-sponsored
6 distortions of competition, (b) not overusing informal settlement processes, as “[w]here the
7 law is ‘grey’, at a certain point agencies have to make an effort to clarify it” (Global Com-
8 petition Review, 2005, 4), (c) being a real enforcer rather than investing mainly in advocacy
9 work, market studies and ‘trial by media’, with courtroom victories being an important indi-
10 cator (Global Competition Review, 2012, 2014), (d) demonstrating leadership in competition
11 policy by means of research and development work and participation in debates on where
12 the lines between pro- and anti-competitive behaviour lie, (e) being a learning organisation,
13 using state of the art methodologies and engaging in continuous self-assessment, (f) taking
14 an economic approach rather than a too formalistic one with an emphasis on the per se rule
15 of anti-competitive behavior, and (g) demonstrating leadership in international co-operation.
16 Hence, this is a rich understanding of quality in terms of process which incorporates substan-
17 tial elements of phronesis, or practical judgement.
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 Because this is ‘found data’, and provided on a commercial basis, we cannot calculate mea-
33 sures of inter-expert reliability. However, GCR states that their star ratings demonstrate con-
34 vergent validity, in that agencies with higher star-ratings were also more likely to be cited by
35 the agencies themselves as ‘most admired’ in a special 2006 survey question (Global Competi-
36 tion Review, 2006, 2). Additionally, the ratings – which are awarded on a scale of one to five,
37 with half-stars permitted (and even quarter stars between 2005 and 2006) – demonstrate a
38 form of test-retest reliability: the average Spearman between agency ratings in adjacent years
39 is high, at 0.958.¹⁰ Finally, the measures correlate reasonably well with two relevant outcome
40 measures from the World Economic Forum’s (WEF) Global Competitiveness Report (Schwab
41 and Sala-i Marin, 2015). The first measure asks survey respondents to rate the “effectiveness
42 of anti-monopoly policy” on a 1-7 scale; the Spearman correlation between GCR rankings and
43 WEF survey responses is moderate ($r = 0.38$). The second measure asks survey respondents
44 to rate the “Extent of market dominance” on a 1-7 scale, reversed so that higher scores indicate
45 lower market dominance. Again, the Spearman correlation between GCR rankings and WEF
46 survey responses is moderate ($r = 0.41$). We use the GCR rankings rather than the WEF mea-
47
48
49
50
51
52
53
54
55
56
57
58
59
60

¹⁰ Note that although we permit quarter-stars in this part of the analysis, for technical reasons relating to the convergence of a probit model estimated with multiple thresholds, we round quarter-marks up to the nearest half-mark in the analysis that follows.

1
2
3 sures because we prefer, for the reasons stated above, using a process measure and because
4
5 the GCR measures are based on individuals (lawyers and economists) who may work for both
6
7 the regulator and the regulatees; the WEF measures, by contrast, are based exclusively on
8
9 business executives, who may have a preference for more lenient competition policy.
10

11 12 13 **Control variables**

14
15 Based on theoretical considerations and the findings of previous studies, we include six control
16
17 variables: size, specialization, experience, agency powers, government effectiveness, and EU
18
19 membership.
20

21
22 First, we include in our models the log of the *number of staff* working on competition policy
23
24 within the agency. This is not equal to the number of staff working within the agency, because
25
26 some staff may work on non-competition related tasks, such as consumer affairs or other ar-
27
28 eas of regulation if the agency is a multi-sectoral regulator. We include this control because
29
30 agencies that are in some sense bigger may produce better work because they can afford to
31
32 specialize in different sectors of the economy or in different types of analysis. We take the log
33
34 of this number because we expect the effects of specialization to show a decreasing marginal
35
36 rate of return. Information on this variable comes from successive editions of the *Global Com-*
37
38 *petition Review*. Missing values for a particular agency-year observation were carried forward
39
40 (backwards) from the last (next) agency-year with non-missing information.
41

42
43 Second, as a more targeted measure of specialization, we include in our models the *propor-*
44
45 *tion of staff* within the agency who work on competition issues (rather than on, for example,
46
47 consumer protection or sectoral regulation). We include this control variable because formal
48
49 modelling suggests that specialization can bolster independence (Dewatripont et al., 1999).
50
51 Information on this variable comes from successive editions of the *Global Competition Review*.
52

53
54 Third, we include the log of (one plus) the *number of years since the establishment* of a competi-
55
56 tion authority organizationally separate from a government ministry. We include this control
57
58 because we believe that agencies with more experience perform better – or that agencies which
59
60 have no previous experience of competition policy to draw upon face teething difficulties in
their early years. We use information from Jordana et al. (2011) to identify the year of estab-
lishment of the first organizationally-separate competition authority.

Fourth, we include a measure of the agency's *powers*. We take into consideration what can be seen as the 'basic powers' of IRAs, with variation across agencies still present. These powers are: (1) the initiation of investigations, (2) the imposition of administrative fines, (3) the introduction of generally-binding rules (delegated or secondary legislation), and (4) the establishment of rules of procedure. The measure is similar in spirit to our measures of independence and accountability, the aggregation being based on an analysis of the responses to the four items. The items loadings are presented in Table 4.

Table 4: *Powers items*

Item	Categories	Location	Discrimination
Can the agency establish its own rules of procedure?	No		3.3
	Yes, and no approval needed	-3.9	3.3
	Yes, and approval needed	-2.5	3.3
Can the agency introduce general binding rules?	No		1.2
	Yes	1.3	1.2
Can the agency initiate investigations?	No		0.7
	Yes	-4.2	0.7
Can the agency impose (administrative) fines on companies?	No		0.2
	Yes	-1.3	0.2

Fifth, we include in our models a measure of *government effectiveness* in each country. We use this measure as a proxy for bureaucratic capacity or "the ability to accomplish intended actions" (Huber and McCarty, 2004, 481). Bureaucratic capacity may be low for a variety of reasons, including the lack of personal capacity of staff, the breakdown and instability of organizational structures, and the presence of incentives for corruption (idem). Importantly, as Huber and McCarty (2004) point out, reductions in capacity not only reduce the general quality of policy-making via straightforward efficiency loss, but also diminish the incentives of civil servants to comply with legislation. Both are relevant for our outcome of interest – the quality of regulatory decision-making. Bureaucratic capacity is a feature of the system as a whole, but it is hypothesized to affect all parts of the bureaucracy, including the competition authority. Hence, we may say that competition authorities 'inherit' a base level of capacity from the rest of the bureaucratic system. To construct the measure, we use data from the World Bank's Worldwide Governance Indicators (Kaufmann et al., 2010). Specifically, we use information concerning the government effectiveness component, which captures

1
2
3 “perceptions of the quality of public services, the quality of the civil service and the
4 degree of its independence from political pressures, the quality of policy formula-
5 tion and implementation, and the credibility of the government’s commitment to
6 such policies” (2010, 4).
7
8
9

10
11
12 The World Bank’s measure is based on fifteen separate sources. We are confident that we
13 can use a measure of the effectiveness of public services to explain the quality of a particular
14 public service, because none of the sources specifically ask about competition policy, and those
15 sources which do ask about effectiveness in specific areas ask about policy areas which are
16 very different to competition policy, such as education, healthcare, transportation and utilities.
17
18 Missing values for a particular agency-year observation were carried forward (backwards)
19 from the last (next) agency-year with non-missing information.
20
21
22
23
24
25

26
27 Sixth, and finally, we include a dummy variable which takes on the value of one if the agency
28 operates in an *EU member state*. We include this variable because EU member states are ex-
29 posed to a form of horizontal accountability through their relationship with DG Competition.
30 Also, they benefit from exchange of best practice through their membership in the European
31 Competition Network (Maggetti and Gilardi, 2014). We expect this to lead to better perfor-
32 mance of competition authorities working in EU member states.
33
34
35
36
37
38
39
40

41 5 Analysis

42
43 In the previous section, we discussed the ratings provided by the GCR. We consider the overall
44 rankings ordinal rather than interval data, such that although four stars is better than three
45 stars, the difference between four and three stars may or may not be the same as the difference
46 between three and two stars. We therefore use a ordinal probit regression model.
47
48
49
50
51

52 In our case, our observations – regulator years – are not independent of one another. Rather,
53 we consider observations as nested within countries, and include random intercepts for each
54 country to model this. Although the observations are ordered, we do not explicitly model
55 the dynamics of these ratings.¹¹ We do, however, include year fixed effects in some models
56
57
58
59
60

¹¹ We do not expect regulatory quality to automatically carry over from year to year over and above the effects of the relevant independent variables. Nonetheless, as a robustness check, we have estimated a model with an autoregressive parameter, which approximates the use of a lagged dependent variable in the ordinal case. The

to allow for the possibility that the GCR became harsher or more lenient in different years. Formally, we posit an underlying latent level of quality, y_i^* , which depends on our predictor variables X and coefficients β , and random effects for countries $c = 1 \dots C$:

$$y_i^* = x_i\beta + \eta_c + \epsilon_i$$

$$\epsilon \sim N(0, \sigma^2)$$

$$\eta_c \sim N(0, \sigma_\eta^2)$$

The latent level of quality is related to the observed rating j ($j = 0 \dots J$) through a series of thresholds:

$$y_i = 0 \Leftrightarrow y_i^* < \tau_1$$

$$y_i = j \Leftrightarrow \tau_j < y_i^* \leq \tau_{j+1}$$

$$y_i = J \Leftrightarrow y_i^* > \tau_J$$

The model is identified by setting σ^2 to one, and by omitting an intercept (Jackman, 2009), and is estimated using the `ordinal` package for R.

Table 5 shows the results of six different models. Models 1 and 2 show just the effects of independence and accountability, net of any controls. Models 3 and 4 include our control variables. Finally, models 5 and 6 include an interaction between independence and accountability. Models 2, 4 and 6 are different from models 1, 3 and 5 in that they include year fixed effects (not shown). These year fixed effects are included to allow for the possibility that the GCR might have become more or less generous with its ratings over time. The values of the coefficients represent the change in the latent level of quality associated with a one-unit increase in the relevant independent variable.

Before discussing the effects of particular variables, we will look at the fit of the different models. The fit of models with year fixed effects is better, but not significantly so: a log-likelihood ratio test shows that Model 4 does not fit significantly better than Model 3 (7.16 on

results – which we report in Table A1 in the Online Appendix – are similar to the ones presented in Table 5. The effect of independence is positive and significant and the effect of accountability is negative, though the latter is not significant in the autoregressive model.

Table 5: Regression model results

	Dependent variable:					
	GCR Rating					
	(1)	(2)	(3)	(4)	(5)	(6)
Independence	0.968*** (0.258)	0.921*** (0.260)	0.757** (0.235)	0.687** (0.241)	0.726** (0.226)	0.658** (0.230)
Accountability	-0.725 (0.399)	-0.823* (0.415)	-0.696* (0.336)	-0.760* (0.340)	-0.749* (0.331)	-0.790* (0.332)
log(No. staff)			0.840*** (0.254)	0.994*** (0.283)	0.821** (0.254)	0.989*** (0.280)
Proportion competition staff			0.344 (0.499)	0.206 (0.511)	0.338 (0.499)	0.186 (0.511)
log(Experience in years)			-0.084 (0.322)	-0.210 (0.401)	-0.089 (0.319)	-0.170 (0.396)
Powers			0.984*** (0.254)	1.016*** (0.255)	0.911*** (0.255)	0.954*** (0.255)
Govt effectiveness			2.033*** (0.470)	2.027*** (0.504)	1.954*** (0.465)	1.937*** (0.495)
EU member-state			1.340 (0.707)	1.354 (0.715)	1.416* (0.684)	1.448* (0.689)
Independence × accountability					0.392 (0.257)	0.349 (0.254)
Year fixed effects?	No	Yes	No	Yes	No	Yes
Observations	315	315	315	315	315	315
Log Likelihood	-341.493	-338.119	-323.873	-320.089	-322.736	-319.180

Note: *p<0.05; **p<0.01; ***p<0.001

1
2
3 10 d.f., $p = 0.71$), and Model 6 does not fit significantly better than Model 5 (6.57 on 10 d.f.,
4
5 $p = 0.77$). There is therefore no evidence of grade inflation (or deflation) over the years.
6
7

8 The fit of the models with an interaction term is also not significantly better than the fit of
9
10 models without an interaction term. For the models without year fixed effects, Model 5 does
11
12 not fit significantly better than Model 3 (2.55 on 1 d.f., $p = 0.11$). Thus, although the coefficient
13
14 has the expected sign, we cannot confirm Hypothesis 3 (that the effects of independence are
15
16 greater given higher levels of accountability).
17

18 We now turn to the interpretation of the findings. In short, the effects of independence are
19
20 positive and highly significant. However, contrary to expectations, a more extensive set of
21
22 accountability provisions is associated with lower rather than higher quality of work, though
23
24 the effect is not always significant. Concerning other agency features, we find a highly signif-
25
26 icant effect of size: bigger agencies are better-rated. The same is true of agencies with powers
27
28 to do more things. Interestingly, we find no effect of experience, nor do we find any bene-
29
30 fit to specialization, contra Dewatripont et al. (1999). Concerning country characteristics, we
31
32 find a highly significant effect of general government effectiveness, and a positive effect of EU
33
34 membership.
35

36 Let us now have a closer look at the findings. Hypothesis 1, concerning formal political
37
38 independence, is confirmed: higher levels of formal political independence are associated with
39
40 higher quality of work.¹² Although our quantitative analysis does not allow us to explore the
41
42 causal mechanism empirically, we expect the effect to work via the increase in expertise that
43
44 associated with higher levels of independence, via the increase in consistency associated with
45
46 the variable, or via both intervening variables. Importantly, our findings suggest that there
47
48 are good reasons to justify political independence by reference to better work.
49

50 The findings on political accountability raise more complicated questions. Hypothesis 2,
51
52 which referred to a positive effect of accountability on quality of work, is not confirmed.
53
54 If anything, there is an effect in the opposite direction, though it is not significant in all model
55
56 specifications.¹³ That is, increases in the type of accountability that we found in the design
57

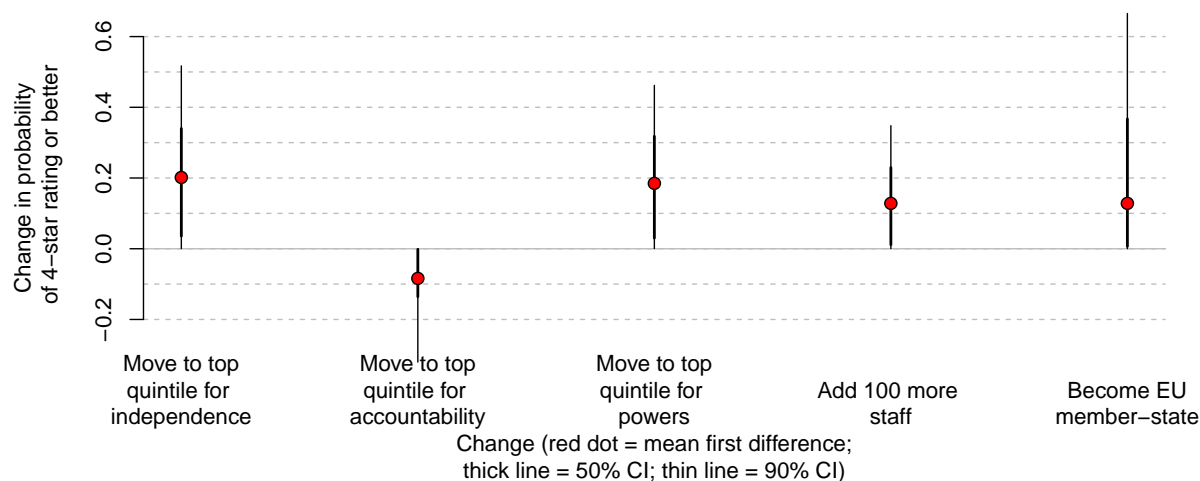
58 ¹² The results are largely the same when we use 'raw' measures of independence (Table A2 in the Online
59 Appendix), except in the models which include the interaction term; in these two models, the effect is neither
60 significant nor positive.

¹³ The effect of accountability is negative and significant in Models 2-5 in Table 5. Yet, when we use 'raw'
measures of independence and accountability (Table A2 in the Online Appendix), accountability ceases to be a

1
2
3 of competition authorities are not associated with higher quality of work. The question is
4 whether this is a more general problem of accountability or a problem specific to the provi-
5 sions we found in the statutes of the agencies. On the one hand, as we pointed out in Section
6 3, accountability provisions introduce reporting procedures which add to the workload of
7 agencies. Indeed, the literature has linked accountability to red tape and excessive costs. On
8 the other hand, previous studies have associated the negative effects with a specific category
9 of provisions: provisions that are generic in nature, including requirements to produce annual
10 plans and reports (Bovens and Schillemans, 2014). These are exactly the kind of provisions
11 that we found in the design of competition authorities. Moreover, we only look at one type
12 of relationship when it comes to accountability: the accountability of competition authorities
13 to politicians. The negative effect of accountability may be relevant only for provisions for
14 political accountability. Even if such provisions come into effect *after* regulatory decisions
15 are taken – thus differing from political independence –, agencies may anticipate the prefer-
16 ences of their political account-holders, leading to the reintroduction of politics via the back
17 door. This is, for instance, why Majone (1999) advocates accountability to other actors than
18 politicians. It might also account for the absence of an interaction effect of independence and
19 accountability. All in all, we need to be careful when it comes to the interpretation of the
20 accountability findings: there are reasons to believe they may be specific to the generic and
21 political accountability that we looked at.¹⁴

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40 The role of EU membership also deserve some more reflection. We included EU membership
41 because previous studies referred to the relevance of authorities' co-ordination with the DG
42 Competition and other competition authorities in the context of the European Competition
43 Network. Such co-ordination is primarily aimed at ensuring that cases are allocated among
44 national competition authorities, and that case-relevant information is being exchanged. This
45 became crucial after the introduction of Regulation 1/2003, which increased the scope of the
46 responsibilities of national competition authorities to also include trade between EU member
47 states; not only trade within their own member state. Yet, while the different actors, including
48 significant predictor. Similarly, when we use the ratio of staff to the size of the economy (Table A3), accountability
49 ceases to be significant at the five percent level.

50
51
52
53
54
55
56
57
58
59
60
¹⁴ In addition, one may argue that accountability provisions may be a response to regulatory quality. That is, governments may increase the accountability of agencies when faced with poor performance. To the extent that there is such a response to performance, we expect it to take an informal form rather than the form of formal provisions. Nonetheless, we have estimated a set of models which explore the potential of such reversed causality (Tables A4-A9 in the Online Appendix). In none of the models is there a significant effect of the quality ratings on accountability, suggesting that there is no convincing evidence that reverse causality can explain the negative association we find, in most models, between accountability and regulatory quality.

Figure 1: *Effect of specified changes on GCR rating*

the DG Competition, do not have any formal power over one another, the co-ordination has resulted in an increase in horizontal accountability, learning from best practices, and even adoption of 'soft law' (Maggetti and Gilardi, 2014). Hence, supranational co-ordination has had an effect of national policy-making (ibid). Although our analysis cannot tell us what the causal mechanism is exactly, it suggests that the co-ordination has not only affected the type of policies introduced at the national level, but also the quality of regulatory decision-making, which has improved.

Because the coefficients in Table 5 represent shifts on a latent scale, and because the latent scale is hardly intuitive, the substantive importance of the coefficients can be put into context by considering their effect on the ratings awarded by the GCR. Here we consider the effects shown in our preferred model, Model 3.¹⁵

Figure 1 shows the effect of specified changes on the probability of receiving a ranking of four stars or more, when all other variables which might affect rankings are held at their mean.¹⁶ When all variables are set to their mean, the probability of receiving a four-star ranking or better is relatively low, at 20%.

When we move independence from the mean to the top quintile, the probability of a four-

¹⁵ We prefer Model 3 because it is the best-performing model according to the Akaike Information Criterion (and thus explains the observed data both well and relatively parsimoniously). The effects under different models would be similar, even for coefficients which are not significant at the 5% level.

¹⁶ These first differences are calculated by sampling from the joint distribution of the coefficients, but taking the threshold parameters as fixed.

1
2
3 star ranking doubles, to just over 40%. The distribution of first differences is skewed, and
4
5 in many simulations the effect of an increase in independence is much greater. When we
6
7 make a comparable change to accountability, by moving it to the top quintile, the probability
8
9 of a four-star ranking decreases, to 11.6%. The magnitude of this change is smaller, despite
10
11 the larger coefficient. This is because the changes in independence move the agency towards
12
13 the steepest part of the curve of the probit function, pushing the agency over the four-star
14
15 threshold, whereas changes in accountability move the agency away from this threshold, and
16
17 make a low-probability event less likely. The effects of a change in powers are more similar to
18
19 the effects of a change in independence.

20
21 Somewhat surprisingly, the effects of these design changes are greater than the effect of adding
22
23 100 more staff to the agency. The distribution of staffing levels is very skewed as a result of
24
25 some very large authorities, which means that a change of 100 staff is actually smaller than a
26
27 change of one standard deviation (155 staff). Nevertheless, this change is a very considerable
28
29 one given that fully half of our observations have fewer than 100 staff, and some agencies get
30
31 by with as few as seventeen competition staff (the case of the Austrian Cartel Office in 2004).¹⁷

32
33 The effect of EU membership is comparable in magnitude to the effect of an increase in inde-
34
35 pendence. In practice, the (true, unknown) effect of EU membership is likely to be greater if
36
37 membership 'causes' countries to increase the level of independence given to their competition
38
39 authority.

40 41 42 43 44 **6 Conclusion and discussion**

45
46 This study has aimed to assess the impact of two much-discussed institutional features – for-
47
48 mal political independence and accountability – on the quality of regulatory decision-making
49
50 in competition authorities. Using the GCR ratings of the performance of competition au-
51
52 thorities in all OECD member states, and controlling for factors such as the number of staff,
53
54 the experience of staff, the powers of the agency, and government effectiveness, our findings
55
56 suggest that independence has a positive and significant effect on the quality of regulatory
57

58 ¹⁷ An alternative view is that resources only matter in relation to the extent of work to be carried out by the
59 competition authority. On this view, what matters is not the level of resources, but their adequacy. We include
60 in the Online Appendix a model which instead features the number of staff divided by the size of the economy
(Table A3). Yet, operationalized this way, staff numbers are no longer a significant predictor of quality. Yet, the
(natural log of the) overall number of staff still is.

1
2
3 policy-making. The same is not true of accountability, which does not boost quality, and
4
5 which does not seem to mediate the effects of independence.
6
7

8 By focusing on the effects of institutional design on the quality of regulatory decision-making,
9
10 this study has moved beyond the existing literature, which has largely concentrated on the
11
12 determinants of institutional design rather than their effects. We have been able to take this
13
14 step thanks to the use of GCR competition authority ratings. This proxy measure allows
15
16 us to capture a wide range of (process-related) elements of the multidimensional concept of
17
18 performance, and is therefore much less subject to the biases resulting from capturing some
19
20 elements of performance rather than others (cf. [Bevan and Hood, 2006](#)), and to the difficulty
21
22 of disentangling the effect of organizational features when using outcome-based measures.
23

24 Our findings bear out the claim of the early literature on delegation to independent agencies.
25
26 Institutions which are more independent tend to do better quality work than institutions
27
28 which are less independent. The effect may be mediated by expertise, policy consistency,
29
30 or both. The question of whether this improved performance sufficiently justifies removing
31
32 competition policy from the scope of normal democratic politics is, of course, a normative
33
34 question. But the presumption held by those in favor of delegation has now received stronger
35
36 empirical backing.
37

38 The findings with respect to accountability are less straightforward to interpret. In our study,
39
40 we made a distinction between independence and accountability. We did so both on the basis
41
42 of conceptual arguments, and on the basis of previous literature which had suggested that
43
44 the designers of independent regulatory agencies could “have their cake and eat it” – that
45
46 they could design for high independence and high accountability. Had we found a positive
47
48 effect of accountability on quality, then regulatory agencies would have been able to live in
49
50 the best of all possible worlds, equipped with high independence, high accountability, and
51
52 producing work of high quality. Instead, our findings point to some trade-off between formal
53
54 political accountability and performance, though the negative effect of accountability is not
55
56 always significant. At the very least, the accountability provisions imposed on competition
57
58 authorities do not seem to help the organizations.
59

60 As discussed in Section 5, the question that our study cannot answer is whether this is a
more general problem of accountability or a problem of accountability design. Scholars do

1
2
3 not unequivocally applaud accountability; they also associate it with red tape and excessive
4 costs, particularly in the case of standardized and routine requirements (Bovens and Schille-
5 mans, 2014). Indeed, the accountability arrangements that we found in the statutes of com-
6 petition authorities fall into the latter category, including standardized provisions on annual
7 budgets, reports, and financial accounts. Moreover, we have solely captured accountability to
8 politicians, while there may be reasons to believe that such accountability reintroduces pol-
9 itics via the back door (cf. Majone, 1999). Our findings may, therefore, support arguments
10 about the sensitivity of accountability design; at least, we have learned that introducing a set
11 of generic requirements for political accountability does not have desirable effects. Future
12 research should pay much more attention to different types of accountability, and potential
13 differential effects. Is there a difference between the effect of generic and more specific ac-
14 countability provisions? How do differences in sanctions linked to accountability matter?
15 And does the effect of political accountability differ from the effect of accountability to other
16 parties? Addressing these types of questions may lead us to find out whether designing
17 accountability more carefully can make a difference.

18
19
20 To the extent that our findings support arguments linking accountability to red tape and
21 excessive costs, they raise the question how desirable accountability really is. In our view, ac-
22 countability serves more purposes than increasing the quality of regulatory decision-making.
23 In particular, political accountability is crucial from a perspective of democratic legitimacy.
24 Regulatory agencies make policies and policy decisions which affect the economy and soci-
25 ety as a whole, and they do so with a fair amount of discretion. It is precisely under these
26 conditions that democratic legitimacy becomes relevant. If we want to ensure that regulatory
27 policies and decisions are still one way or another linked to citizens, we may want to introduce
28 accountability provisions even if they come at a cost. Nonetheless, if the effect of accountabil-
29 ity on regulatory quality is dependent on the form that it takes, there is all reason to invest in
30 analysing and improving the design.

31
32
33 We also stressed that our findings apply to de jure rather than de facto independence and ac-
34 countability. Agencies may have low de jure accountability, in that they are not compelled by
35 legislation to report back on their activities, but may have high de facto accountability, in that
36 they provide such information of their own initiative. It is not clear from our work whether
37 such 'voluntary accountability' (Koop, 2014) on the part of the agency would harm perfor-

1
2
3 mance. Future research into the links between de facto independence, accountability and
4 quality will require in-depth examination of agencies' efforts to make themselves accountable,
5 and in particular whether agencies' accountability obligations are discharged in a formalistic
6 manner, or whether instead they provide an insight into the decision-making process of the
7 agency and, thereby, foster greater predictability and certainty for market operators.
8
9

10
11
12
13 Finally, we started our research with an interest in the quality of decision-making by inde-
14 pendent regulatory agencies. Our analysis raises the question to what extent the findings can
15 'travel' to other types of regulatory agencies. In our view, there is no reason to think that the
16 findings are competition authority-specific. The arguments linking political independence to
17 better performance are relevant for all regulatory agencies; not just for competition authori-
18 ties. Equally, all arguments related to accountability – in the theoretical section as well in the
19 discussion of the results – are more generic. There is one key difference, though: while com-
20 petition authorities regulate (specific aspects of) the economy as a whole, most other agencies
21 regulate specific sectors. This increases the importance of analysing the independence of the
22 regulator from the regulatory sector. While regulators can benefit from interacting with their
23 regulatees – particularly, in terms of information and expertise –, they may also be captured
24 by them, leading to a decrease in the quality of decision-making. As such capture is much
25 more likely in sectoral regulation, independence from the sector may need to be taken into
26 account in studies of the quality of decision-making by sectoral regulators.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 **Acknowledgements**

43
44 We have greatly benefited from the comments and suggestions offered by seminar participants
45 at the University of Exeter, King's College London, the Centre for Competition Policy at the
46 University of East Anglia, the 2014 conference of the ECPR Standing Group on Regulatory
47 Governance, and the 2014 conference on 'Consequences of Multilevel Governance' held at the
48 Hanse-Wissenschaftskolleg in Delmenhorst. We would also like to thank Antje Kreutzman-
49 Gallasch for help with gathering data on regulators' statutes, and the Centre for Competition
50 Policy at the University of East Anglia for funding the data collection. Finally, we would like
51 to thank the anonymous reviewers, whose comments strengthened the paper a great deal.
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6 Bawn, K. (1995). Political control versus expertise: Congressional choices about administrative proce-
7 dures. *American Political Science Review*, 89(1):62–73.
8
9
10 Bernstein, M. H. (1955). *Regulating Business by Independent Commission*. Princeton, NJ: Princeton Uni-
11 versity Press.
12
13
14 Bevan, G. and Hood, C. (2006). What’s measured is what matters: Targets and gaming in the English
15 public health care system. *Public Administration*, 84(3):517–538.
16
17
18 Bortolotti, B., Cambini, C., Rondi, L., and Spiegel, Y. (2011). Capital structure and regulation: Do own-
19 ership and regulatory independence matter? *Journal of Economics & Management Strategy*, 20(2):517–
20 564.
21
22
23
24 Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law*
25 *Journal*, 13(4):447–468.
26
27
28 Bovens, M. and Schillemans, T. (2014). Meaningful accountability. In Bovens, M., Goodin, R. E., and
29 Schillemans, T., editors, *The Oxford Handbook of Public Accountability*, pages 673–682. Oxford: Oxford
30 University Press.
31
32
33
34 Busuioc, M. (2009). Accountability, control and independence: The case of European agencies. *European*
35 *Law Journal*, 15(5):599–615.
36
37
38 Cukierman, A., Webb, S. B., and Neyapti, B. (1992). Measuring the independence of central banks and
39 its effect on policy outcomes. *World Bank Economic Review*, 6(3):353–398.
40
41
42 Dewatripont, M., Jewitt, I., and Tirole, J. (1999). The economics of career concerns, part II: Application
43 to missions and accountability of government agencies. *Review of Economic Studies*, 66(1):199–217.
44
45
46 Dubnick, M. J. and Frederickson, H. (2011). Introduction. In Dubnick, M. J. and Frederickson, H. G.,
47 editors, *Accountable Governance: Problems and Promises*, pages xi–xxx. Armonk, NY: M.E. Sharpe.
48
49
50 Edwards, G. and Waverman, L. (2006). The effects of public ownership and regulatory independence
51 on regulatory outcomes. *Journal of Regulatory Economics*, 29(1):23–67.
52
53
54 Elgie, R. and McMenamin, I. (2005). Credible commitments, political uncertainty or political com-
55 plexity? Explaining variation in the independence of non-majoritarian institutions in France. *British*
56 *Journal of Political Science*, 35(3):531–548.
57
58
59
60 Ennsner-Jedenastik, L. (2015). Credibility versus control: Agency independence and partisan influence
in the regulatory state. *Comparative Political Studies*, 48(7):823–853.

- 1
2
3 Gailmard, S. and Patty, J. W. (2007). Slackers and zealots: Civil service, policy discretion, and bureau-
4 cratic expertise. *American Journal of Political Science*, 51(4):873–889.
5
6
7 Gilardi, F. (2002). Policy credibility and delegation to independent regulatory agencies: A comparative
8 empirical analysis. *Journal of European Public Policy*, 9(6):873–893.
9
10
11 Gilardi, F. (2005). The institutional foundations of regulatory capitalism: The diffusion of independent
12 regulatory agencies in Western Europe. *Annals of the American Academy of Political and Social Science*,
13 598:84–101.
14
15
16
17 Global Competition Review (2005). *Rating Enforcement 2005*. London: Law Business Research.
18
19
20 Global Competition Review (2006). *Rating Enforcement 2006*. London: Law Business Research.
21
22
23 Global Competition Review (2007). *Rating Enforcement 2007*. London: Law Business Research.
24
25
26 Global Competition Review (2008). *Rating Enforcement 2008*. London: Law Business Research.
27
28
29 Global Competition Review (2011). *Rating Enforcement 2011*. London: Law Business Research.
30
31
32 Global Competition Review (2012). *Rating Enforcement 2012*. London: Law Business Research.
33
34
35 Global Competition Review (2014). *Rating Enforcement 2014*. London: Law Business Research.
36
37
38 Global Competition Review (2015). *Rating Enforcement 2015*. London: Law Business Research.
39
40
41 Gutiérrez, L. H. (2003). The effect of endogenous regulation on telecommunications expansion and
42 efficiency in Latin America. *Journal of Regulatory Economics*, 23(3):257–286.
43
44
45 Hanretty, C. and Koop, C. (2012). Measuring the formal independence of regulatory agencies. *Journal*
46 *of European Public Policy*, 19(2):198–216.
47
48
49 Hanretty, C. and Koop, C. (2013). Shall the law set them free? The formal and actual independence of
50 regulatory agencies. *Regulation & Governance*, 7(2):195–214.
51
52
53 Hood, C. and Bevan, G. (2006). What’s measured is what matters: Targets and gaming in the English
54 public health care system. *Public Administration*, 84(3):517–538.
55
56
57 Huber, J. D. and McCarty, N. (2004). Bureaucratic capacity, delegation, and political reform. *American*
58 *Political Science Review*, 98(3):481–494.
59
60
61 Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. Chichester, UK: John Wiley & Sons.
62
63
64 Jordana, J., Levi-Faur, D., and Fernández i Marín, X. (2011). The global diffusion of regulatory agencies:
65 Channels of transfer and stages of diffusion. *Comparative Political Studies*, 44(10):1343–1369.

- 1
2
3 Jordana, J. and Rosas, G. (2014). When do autonomous banking regulators promote stability? *European*
4 *Journal of Political Research*, 53(4):672–691.
5
6
7 Kaufmann, D., Kraay, A., and Mastruzzi, M. (2010). The worldwide governance indicators: Methodol-
8 ogy and analytical issues. World Bank Policy Research Working Paper No 5430.
9
10
11 Koop, C. (2011). Explaining the accountability of independent agencies: The importance of political
12 salience. *Journal of Public Policy*, 31(2):209–234.
13
14
15 Koop, C. (2014). Theorizing and explaining voluntary accountability. *Public Administration*, 92(3):565–
16 581.
17
18
19 Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal
20 plans. *Journal of Political Economy*, 85(3):473–492.
21
22
23 Lerner, J. S. and Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*,
24 125(2):255.
25
26
27 Levy, B. and Spiller, P. T. (1996). *Regulations, Institutions, and Commitment: Comparative Studies of Telecom-*
28 *munications*. Cambridge: Cambridge University Press.
29
30
31 Loewenstein, K. (1957). *Political Power and the Governmental Process*. Chicago: University of Chicago
32 Press.
33
34
35 Maggetti, M. and Gilardi, F. (2014). Network governance and the domestic adoption of soft rules.
36 *Journal of European Public Policy*, 21(9):1293–1310.
37
38
39 Maggetti, M., Ingold, K., and Varone, F. (2013). Having your cake and eating it, too: Can regulatory
40 agencies be both independent and accountable? *Swiss Political Science Review*, 19(1):1–25.
41
42
43 Majone, G. (1996). Temporal consistency and policy credibility: Why democracies need non-
44 majoritarian institutions. EUI Working Paper, RSC No 96/57.
45
46
47 Majone, G. (1997). From the positive to the regulatory state: Causes and consequences of changes in
48 the mode of governance. *Journal of Public Policy*, 17(2):139–167.
49
50
51 Majone, G. (1999). The regulatory state and its legitimacy problems. *West European Politics*, 22(1):1–24.
52
53
54 May, P. J. (2007). Regulatory regimes and accountability. *Regulation & Governance*, 1(1):8–26.
55
56
57 McCubbins, M. D., Noll, R., and Weingast, B. (1987). Administrative procedures as instruments of
58 political control. *Journal of Law, Economics, and Organization*, 3(2):243–277.
59
60
61 Mero, N. P. and Motowidlo, S. J. (1995). Effects of rater accountability on the accuracy and the favora-
62 bility of performance ratings. *Journal of Applied Psychology*, 80(4):517–524.

- 1
2
3 Moe, T. M. (2012). Delegation, control, and the study of public bureaucracy. In Gibbons, R. and Roberts,
4 J., editors, *The Handbook of Organizational Economics*, pages 1148–1182. Princeton, NJ: Princeton Uni-
5 versity Press.
6
7
8
9 Motta, M. (2004). *Competition Policy: Theory and Practice*. Cambridge: Cambridge University Press.
10
11 Patil, S. V., Vieider, F., and Tetlock, P. E. (2014). Process versus outcome accountability. In Bovens, M.,
12 Goodin, R. E., and Schillemans, T., editors, *The Oxford Handbook of Public Accountability*, pages 69–89.
13 Oxford: Oxford University Press.
14
15
16
17 Philp, M. (2009). Delimiting democratic accountability. *Political Studies*, 57(1):28–53.
18
19
20 Quintyn, M. and Taylor, M. W. (2007). Robust regulators and their political masters: Independence and
21 accountability in theory. In Masciandaro, D. and Quintyn, M., editors, *Designing Financial Supervision*
22 *Institutions: Independence, Accountability and Governance*, pages 3–40. Cheltenham, Edward Elgar.
23
24
25 Radaelli, C. (2004). Getting to grips with quality in the diffusion of regulatory impact assessment in
26 Europe. *Public Money and Management*, 24(5):271–276.
27
28
29 Raju, N. S., Laffitte, L. J., and Byrne, B. M. (2002). Measurement equivalence: A comparison of meth-
30 ods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*,
31 87(3):517–529.
32
33
34 Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *Quarterly*
35 *Journal of Economics*, 100(4):1169–1189.
36
37
38 Samejima, F. (1970). Estimation of latent ability using a response pattern of graded scores. *Psychome-*
39 *trika*, 35(1):139–139.
40
41
42 Schwab, K. and Sala-i Marin, X. (2015). The global competitiveness report, 2015–2016.
43
44
45 Scott, C. (2000). Accountability in the regulatory state. *Journal of Law and Society*, 27(1):38–60.
46
47
48 Selin, J. (2015). What makes an agency independent? *American Journal of Political Science*, 59(4):971–987.
49
50
51 Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quar-*
52 *terly*, 46(4):285–292.
53
54
55 Tetlock, P. E. (1985). Accountability: A social check on the fundamental attribution error. *Social Psy-*
56 *chology Quarterly*, 48(3):227–236.
57
58
59 Thatcher, M. (2005). The third force? Independent regulatory agencies and elected politicians in
60 Europe. *Governance*, 18(3):347–373.

1
2
3 Thatcher, M. and Stone Sweet, A. (2002). Theory and practice of delegation to non-majoritarian insti-
4 tutions. *West European Politics*, 25(1):1–22.
5
6

7 Wallsten, S. J. (2001). An econometric analysis of telecom competition, privatization, and regulation in
8 Africa and Latin America. *Journal of Industrial Economics*, 49(1):1–19.
9
10

11 Weldon, E. and Gargano, G. M. (1988). Cognitive loafing the effects of accountability and shared
12 responsibility on cognitive effort. *Personality and Social Psychology Bulletin*, 14(1):159–171.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Political Independence, Accountability, and the Quality of Regulatory Decision-Making

Online Appendix*

This online appendix provides details of several additional models which serve as robustness checks on the findings presented in the main article. One section deals with the question of changes over time. Two sections deal with alternative operationalizations of key variables. A final section deals with the possibility of reverse causality between accountability and quality.

1 Changes over time

In the main article we argued that a model with country random effects and year fixed effects was adequate for modelling data from an unbalanced panel with a short number of time periods. It is possible to model dynamics more explicitly. Here, we show that changes in independence and accountability matter for ratings. We do this in a very simple way, before going on to show the same thing in a more complicated way.

To show that changes matter in a simple way, we take each time an agency had its statutes changed, calculate the change in independence (and the change in accountability), and the mean rating before the change, and the mean rating after the change. If an agency changed twice, we take the mean rating after the first change but before the second change. On a number of occasions, there is no change in independence (but there may be a change in accountability, or in powers).

The left hand panel of Figure A1 shows the relationship between changes in independence and changes in ratings; the right hand panel shows the relationship for changes in accountability. The independence relationship is positive and significant at the 10% level ($r = 0.46$ on 13 d.f., $p = 0.09$); the accountability relationship is not ($r = 0$ on 13 d.f., $p = 0.99$).

These bivariate tests on changes (a) “knock out” constant effects of factors which vary across countries but are constant within countries; and (b) ignore many other factors which vary within countries. In order to take account of these additional factors, it is necessary to consider a more complex model.

A more complex model

We argue that our models of quality – which relied on country random intercepts – are appropriate for our data, characterised by a moderate number of cross-sectional units and a limited and varying number of time periods. However, it is possible to include a model which models dynamics more explicitly. The model is a version of the model described in (Stegmueller 2013).

*. Contact:

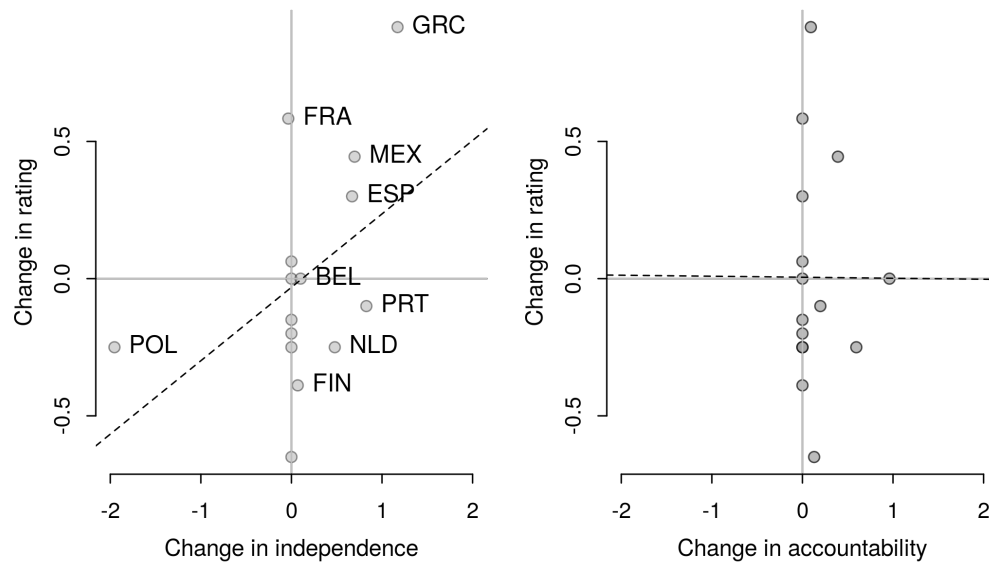


Figure A1: Relationship between changes in independence, accountability and changes in ratings

As before, we posit an unobserved latent level of quality, y^* . We now subscript this latent level of quality using both an index variable for the regulator ($i = 1, \dots, I$) and time ($t = 1, \dots, T$). For periods after the initial period, we model the latent level of quality as a function of country random effects, a matrix of predictor variables, and the previous latent level of quality multiplied by a parameter (ϕ) gauging the extent of autocorrelation.

$$y_{it}^* = \phi y_{it-1}^* + \beta' x_{it} + \eta_i + \epsilon_{it}$$

Where autocorrelation is high, ϕ will be close to one. This equation explains changes with respect to the previous value of y^* . The initial value of y^* is explain by a separate equation:

$$y_{i1}^* = \delta' w_i + \lambda \eta_i + \epsilon_{i1}$$

Here, w_i is a matrix of variables explaining the values of the initial observation, and δ a vector of the coefficients associated with these variables. The random intercepts from the dynamic model are included, but are included using a scale factor (λ) to allow for “a different effect magnitude of unobserved characteristics on initial [quality]”.

Here, we include the design variables (independence, accountability, powers) in the matrix used in the equation modelling changes in levels of quality. We include all variables in the matrix used in the equation modelling initial levels of quality. The results of the model, estimated using Bayesian methods, are presented in Table A1.

As the model shows, independence is significantly associated with positive changes in latent quality. Accountability has a negative effect, but since the 95% credible interval encompasses zero, is it not significant at the 5% level (though it is significant at the 10% level). Changes in powers do not seem to have an effect on changes in latent quality.

Table A1: Models with autoregressive parameter

Variable	Mean	95 pct. CI
Independence	0.210	(0.048, 0.393)
Accountability	-0.254	(-0.541, 0.043)
Powers	-0.068	(-0.397, 0.316)
Intercept [initial conditions]	-19.959	(-28.293, -11.933)
Independence [initial conditions]	0.365	(-0.558, 1.341)
Accountability [initial conditions]	0.325	(-1.068, 1.839)
log(Staff) [initial conditions]	2.093	(0.669, 3.365)
Proportion competition staff [initial conditions]	-2.999	(-6.361, 0.414)
Powers [initial conditions]	2.742	(1.273, 4.307)
Government effectiveness [initial conditions]	6.819	(5.234, 8.598)
EU membership [initial conditions]	4.284	(2.355, 6.308)
phi [degree of autocorrelation]	0.967	(0.916, 0.997)

Surprisingly, independence and accountability do not have significant effects on initial levels of quality, though the extent of the agency's powers does have a significant positive effect. Note, however, that the model explaining initial conditions is trying to predict just thirty-one values using eight estimated parameters, some of which (government effectiveness, EU membership) have very large effects.

The degree of auto-correlation is high, which suggests that initial conditions explain most of what makes a competition authority good.

Note, however, that we lack a theory as to which of the canvassed variables should affect initial conditions, and which should affect dynamics. In the above model, we have included only the design variables (independence, accountability and power) which are the focus of our article. However, it would also be possible to include other variables which change over time, from variables that change relatively quickly (staffing numbers) to those which change relatively slowly (governmental effectiveness). We could include all of the variables, but then we would be estimating around fifty parameters on just over three hundred data points. In the absence of such a theory, we use in the main article a model which is appropriate for the data, and allows us to test a number of cross-sectional and within-unit relationships.

2 Models with raw scores

Table A2 shows the same models as appear in the main article, except that our latent trait measures of independence, accountability and powers have been replaced with raw score measures of independence, accountability and powers. That is, the responses have been scored equally on the interval zero to one, and the simple average of these responses gives the score.

The table shows that independence is positively associated with quality across all models, but that this effect is only significant at the 5% level in models (1), (2) and (3). Independence is significant at the 10% level in model (4). The interpretation of the effects of independence in models (5) and (6) is complicated by the presence of the interaction term.

The effect of accountability is negative across all specifications, but is in no model significant at the 5% or 10% levels.

Table A2: Regression model results, raw measures of independence

	<i>Dependent variable:</i>					
	GCR Rating					
	(1)	(2)	(3)	(4)	(5)	(6)
Independence (raw score)	4.490** (1.716)	3.856* (1.764)	3.630* (1.710)	2.978 (1.714)	-0.192 (4.935)	-0.250 (4.836)
Accountability (raw score)	-2.137 (1.827)	-2.347 (1.862)	-1.486 (1.554)	-1.691 (1.523)	-6.342 (6.169)	-5.878 (6.136)
log(No. staff)			0.622* (0.249)	0.786** (0.270)	0.623* (0.247)	0.783** (0.267)
Proportion competition staff			0.713 (0.492)	0.563 (0.501)	0.748 (0.495)	0.595 (0.503)
log(Experience (years))			0.047 (0.318)	0.015 (0.394)	-0.015 (0.324)	-0.041 (0.398)
Powers (raw score)			2.149 (1.210)	2.368 (1.217)	2.290 (1.208)	2.472* (1.214)
Govt effectiveness			1.764*** (0.476)	1.712*** (0.492)	1.696*** (0.476)	1.658*** (0.492)
EU member-state			0.854 (0.712)	0.945 (0.708)	0.813 (0.692)	0.908 (0.695)
Independence × accountability					7.712 (9.464)	6.647 (9.415)
Year FEs?	No	Yes	No	Yes	No	Yes
Observations	315	315	315	315	315	315
Log Likelihood	-346.545	-343.435	-335.629	-331.999	-335.305	-331.754

Note:

*p<0.05; **p<0.01; ***p<0.001

3 Models with staff divided by GDP

Table A3 shows a series of models which tackle the issue of staffing adequacy, rather than the measure of staffing used in the article (the natural log of the number of staff). All of the models include a variable equalling the number of staff divided by the gross domestic product of the country, measured in trillions of 2005 PPP-adjusted US dollars (Feenstra, Inklaar, and Timmer 2015; Zeileis 2015). Note that we have dropped the year fixed effects from these models.

The rationale for this new variable is as follows. Competition authorities display regulatory quality in their work. However, the amount of work they must or ought to deal with varies depending on the size of the economy. As a result, staff levels which might be sufficient to secure high quality in a small economy might be insufficient in a larger economy. Thus, a measure which takes account of the *adequacy of resources* is preferable to a measure which focuses just on their quantity. We thank the reviewers for this suggestion.

Model (1) show that this variable, on its own, is not significant. However, there are reasons to think that including this variable on its own in the model specification is misguided. It is now well know that interaction terms in a regression model must include all constituent terms. A model without constituent terms assumes (implicitly and often incorrectly) that the constituent terms have zero effect on their own. In our case, the same reasoning applies to dividing staff by GDP: we might think that GDP has an effect on its own.

Model (2) and (3) therefore introduces the constituent terms (GDP in trillions and the raw number of staff). GDP has a significant and positive effect, but neither the raw number of staff nor the measure of the adequacy of staffing are significant.

Model (4) replaces the measure of staffing and the measure of staffing adequacy with measures based on the log of the number of staffing. Here, the measure of staffing adequacy is based on the log number of staff divided by GDP in trillions. Here we see that staffing now does have a significant and positive effect on GCR ratings. The measure of staffing adequacy is significant, but has the sign opposite to that expected: the more “adequate” the staffing, the worse the performance. The better performance of a model which includes the natural log of the number of staff is what we would expect if there was a decreasing marginal benefit of additional staff members.

4 Reverse causality

We present in Tables A4-9 a series of regression models designed to test for reverse causality. The dependent variable in each regression model is the change in accountability between two years.

The key independent variable changes across different model specifications. One set of models tests whether changes in accountability respond to the *level* of the ratings. These levels are lagged by one, two and three years in different specifications (with consequences for the number of observations featured in the model). A further set of models tests whether changes in accountability respond to *changes* in ratings. These changes are also lagged by one, two and three years. The models include many of the control variables featured in our models of qual-

Table A3: Regression model results, staffing adequacy

	<i>Dependent variable:</i>			
	GCR Rating			
	(1)	(2)	(3)	(4)
Independence	1.624*** (0.456)	1.199** (0.433)	1.024* (0.426)	0.794*** (0.002)
Accountability	-1.095 (0.702)	-1.304* (0.631)	-1.233* (0.605)	-1.099*** (0.002)
Staff / GDP (trillions USD)	0.0001 (0.002)	0.001 (0.002)	-0.0004 (0.002)	
log(Staff) / GDP (trillions USD)				-0.056*** (0.002)
Proportion competition staff	1.422 (0.919)	1.159 (0.914)	0.909 (0.927)	0.559*** (0.002)
log(Staff)				1.293*** (0.002)
log(Experience (years))	0.555 (0.593)	0.179 (0.571)	0.086 (0.562)	-0.232*** (0.002)
Powers	1.349** (0.457)	1.329** (0.439)	1.567*** (0.451)	1.695*** (0.002)
Govt effectiveness	2.791** (0.929)	2.742** (0.859)	2.883*** (0.834)	3.465*** (0.002)
EU member-state	2.185 (1.574)	3.195* (1.388)	3.201* (1.300)	2.813*** (0.002)
GDP (trillions USD)		1.009*** (0.286)	0.783** (0.285)	0.640*** (0.002)
Staff (raw count)			0.005 (0.003)	
Observations	315	315	315	315
Log Likelihood	-325.795	-320.106	-318.338	-316.276

Note:

*p<0.05; **p<0.01; ***p<0.001

ity. They also feature country random effects, and year fixed effects where indicated.

In none of the models is there a significant effect of the level of the ratings, or changes in ratings, on levels of accountability. The coefficient generally has the expected sign (negative, such that higher ratings imply fewer accountability demands) in the models using levels, but the sign of the coefficient is mixed in the models using changes in ratings. On these measures, and using these model specifications, there is no convincing evidence that reverse causality can explain the negative association we find between accountability and regulatory quality.

Table A4: Models on levels, one-year lag

	<i>Dependent variable:</i>			
	Accountability.chg			
	(1)	(2)	(3)	(4)
Rating.l	−0.015*	−0.018**	−0.046***	−0.046***
	(0.009)	(0.009)	(0.013)	(0.012)
log(Staff)			0.033***	0.031***
			(0.011)	(0.011)
Proportion competition staff			−0.077***	−0.064**
			(0.029)	(0.028)
log(Experience)			−0.015	−0.018
			(0.013)	(0.012)
Govt effectiveness			0.031**	0.032**
			(0.015)	(0.015)
EU			0.015	0.013
			(0.018)	(0.017)
Constant	0.062**	0.340***	0.065	0.325***
	(0.031)	(0.106)	(0.066)	(0.116)
Year effects	No	Yes	No	Yes
Observations	253	253	253	253
Log Likelihood	215.953	200.249	206.828	190.618
Akaike Inf. Crit.	−423.906	−374.498	−395.656	−345.235
Bayesian Inf. Crit.	−409.773	−328.564	−363.855	−281.634

Note:

*p<0.1; **p<0.05; ***p<0.01

Table A5: Models on levels, two-year lag

	<i>Dependent variable:</i>			
	Accountability.chg			
	(1)	(2)	(3)	(4)
Rating.l2	−0.019** (0.009)	−0.017* (0.009)	−0.049*** (0.012)	−0.050*** (0.013)
log(Staff)			0.032*** (0.010)	0.033*** (0.011)
Proportion competition staff			−0.068** (0.028)	−0.076*** (0.029)
log(Experience)			−0.015 (0.012)	−0.019 (0.013)
Govt effectiveness			0.030** (0.014)	0.035** (0.015)
EU			0.013 (0.017)	0.013 (0.017)
Constant	0.072** (0.031)	−0.051 (0.082)	0.074 (0.061)	−0.049 (0.095)
Year effects	No	Yes	No	Yes
Observations	255	255	255	255
Log Likelihood	218.921	200.709	209.672	191.714
Akaike Inf. Crit.	−429.842	−375.418	−401.343	−347.427
Bayesian Inf. Crit.	−415.677	−329.381	−369.472	−283.685

Note:

*p<0.1; **p<0.05; ***p<0.01

Table A6: Models on levels, three-year lag

	<i>Dependent variable:</i>			
	Accountability.chg			
	(1)	(2)	(3)	(4)
Rating.l3	−0.002 (0.010)	0.002 (0.010)	−0.020 (0.015)	−0.012 (0.015)
log(Staff)			0.019 (0.012)	0.014 (0.012)
Proportion competition staff			−0.056* (0.033)	−0.056* (0.033)
log(Experience)			−0.017 (0.014)	−0.021 (0.014)
Govt effectiveness			0.016 (0.017)	0.017 (0.017)
EU			0.004 (0.020)	−0.0004 (0.020)
Constant	0.019 (0.035)	−0.269** (0.117)	0.060 (0.073)	−0.202 (0.129)
Year effects	No	Yes	No	Yes
Observations	225	225	225	225
Log Likelihood	176.635	161.797	163.395	148.436
Akaike Inf. Crit.	−345.270	−297.595	−308.789	−260.871
Bayesian Inf. Crit.	−331.605	−253.186	−278.045	−199.382

Note:

*p<0.1; **p<0.05; ***p<0.01

Table A7: Models on changes, one-year lag

	<i>Dependent variable:</i>			
	Accountability.chg			
	(1)	(2)	(3)	(4)
Rating change.l	−0.009 (0.026)	−0.014 (0.026)	−0.004 (0.026)	−0.009 (0.027)
log(Staff)			0.008 (0.009)	0.007 (0.009)
Proportion competition staff			−0.050* (0.029)	−0.042 (0.028)
log(Experience)			−0.013 (0.012)	−0.016 (0.012)
Govt effectiveness			0.007 (0.013)	0.010 (0.013)
EU			0.0001 (0.017)	−0.002 (0.016)
Constant	0.011* (0.006)	0.261*** (0.099)	0.042 (0.064)	0.284** (0.119)
Year effects	No	Yes	No	Yes
Observations	253	253	253	253
Log Likelihood	215.656	199.484	201.293	184.659
Akaike Inf. Crit.	−423.311	−372.968	−384.586	−333.319
Bayesian Inf. Crit.	−409.178	−327.034	−352.786	−269.718

Note:

*p<0.1; **p<0.05; ***p<0.01

Table A8: Models on changes, two-year lag

	<i>Dependent variable:</i>			
	Accountability.chg			
	(1)	(2)	(3)	(4)
Rating change.l	−0.008 (0.022)	−0.004 (0.023)	−0.004 (0.022)	−0.0004 (0.023)
log(Staff)			0.007 (0.009)	0.006 (0.009)
Proportion competition staff			−0.046 (0.029)	−0.044 (0.029)
log(Experience)			−0.013 (0.012)	−0.017 (0.012)
Govt effectiveness			0.007 (0.013)	0.010 (0.013)
EU			−0.002 (0.017)	−0.003 (0.017)
Constant	0.011* (0.006)	0.0004 (0.020)	0.045 (0.064)	0.046 (0.065)
Year effects	No	Yes	No	Yes
Observations	251	251	251	251
Log Likelihood	215.872	197.840	201.040	183.103
Akaike Inf. Crit.	−423.744	−371.681	−384.079	−332.206
Bayesian Inf. Crit.	−409.643	−329.375	−352.350	−272.273

Note:

*p<0.1; **p<0.05; ***p<0.01

Table A9: Models on changes, three-year lag

	<i>Dependent variable:</i>			
	Accountability.chg			
	(1)	(2)	(3)	(4)
Rating change.l3	-0.086*** (0.021)	-0.101*** (0.021)	-0.085*** (0.021)	-0.100*** (0.021)
log(Staff)			0.009 (0.010)	0.007 (0.009)
Proportion competition staff			-0.036 (0.031)	-0.043 (0.030)
log(Experience)			-0.016 (0.014)	-0.019 (0.014)
Govt effectiveness			0.003 (0.015)	0.008 (0.015)
EU			-0.002 (0.018)	-0.003 (0.018)
Constant	0.011* (0.007)	-0.312*** (0.101)	0.046 (0.070)	-0.271** (0.120)
Year effects	No	Yes	No	Yes
Observations	225	225	225	225
Log Likelihood	185.682	173.955	170.839	159.517
Akaike Inf. Crit.	-363.365	-321.909	-323.679	-283.033
Bayesian Inf. Crit.	-349.701	-277.500	-292.934	-221.543

Note:

*p<0.1; **p<0.05; ***p<0.01

References

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The next generation of the Penn world table." *American Economic Review* 105(10): 3150–3182. <http://www.ggd.c.net/pwt/>.

Stegmueller, Daniel. 2013. "Modeling dynamic preferences: A Bayesian robust dynamic latent ordered probit model." *Political Analysis* 21(3): 314–333.

Zeileis, Achim. 2015. *Pwt8: Penn world table (version 8.0)*. <http://CRAN.R-project.org/package=pwt8>.