

THE IMPORTANCE OF INPUT AND OUTPUT LEGITIMACY IN DEMOCRATIC  
GOVERNANCE. EVIDENCE FROM A POPULATION-BASED SURVEY EXPERIMENT  
IN FOUR WEST EUROPEAN COUNTRIES

Michael Andrea Strebel\*  
University of Zurich

Daniel Kübler  
University of Zurich

Frank Marcinkowski  
Heinrich Heine  
University Düsseldorf

\*Corresponding Author:  
[michael.strebel@uzh.ch](mailto:michael.strebel@uzh.ch)

Paper accepted for publication in the *European Journal of Political Research*

ABSTRACT

The study of subjective democratic legitimacy from a citizens' perspective has become an important strand of research in political science. Echoing the well-known distinction between 'input-' and 'output-oriented' legitimacy, the scientific debate on this topic has coined two opposed views. Some scholars find that citizens have a strong and intrinsic preference for meaningful participation in collective decision-making. But others argue, to the contrary, that citizens prefer 'stealth democracy' because they care mainly about the substance of decisions, but much less about the procedures leading to them.

In this article, we explore citizens' preferences regarding democratic governance, focusing on their evaluations of a public policy according to criteria related to various legitimacy dimensions, as well as on the (tense) relationship among them. We do so using data from a population-based conjoint experiment conducted in eight metropolitan areas in France, Germany, Switzerland, and the United Kingdom. By analyzing 5'000 respondents' preferences for different governance arrangements, that were randomly varied with respect to

their input-, throughput- and output-quality as well as their scope of authority, we shed a light on the relative importance of different aspects of democratic governance.

We find, first, that output evaluations are the most important driver for citizens' choice of a governance arrangement. Second, we find consistent positive effects of criteria of input- and throughput-legitimacy that operate largely independent of output evaluations. Finally, democratic input, but not democratic throughput, is considered somewhat more important when a governance body holds a high level of formal authority.

These findings run counter to a central tenet of the 'stealth democracy' argument. While they indeed suggest that political actors and institutions can gain legitimacy primarily through the provision of 'good output', citizens' demand for input and throughput do not seem to be conditioned by the quality of output as advocates of stealth democratic theory suggest. Democratic input and throughput remain important secondary features of democratic governance.

**KEYWORDS:** Democratic Legitimacy, Democratic Governance, Stealth Democracy, Public Opinion, Conjoint Analysis

## INTRODUCTION

What do citizens' expect from democratic governance? What are the crucial factors citizens rely on to assess a political system and its outputs? Political scientists have been studying these issues for decades (Easton 1965; Norris 1999). These questions are relevant because they engage with the core of a political order's legitimacy. In democracies, the public policy process - which we can define as coordinated action aimed at resolving public problems - must be coherent with normative principles shared by the citizens. Only when this is the case, conventional wisdom assumes, will people be likely to accept outcomes even when they personally think that these are not desirable.

What are the normative principles of democratic governance that citizens cherish? The answer to this question is currently subject to debate. Existing research has crystallized two prominent but diametrically opposed views on citizens' preferences concerning their inclusion in political decision-making processes. Some scholars find that citizens strongly prefer to participate in collective decision-making (e.g. Tyler 2000, Bengtsson and Mattila 2009, Esaiasson et al. 2012). But findings by others suggest the opposite, namely that citizens have strong preferences for "stealth democracy" (Hibbing and Theiss-Morse 2002), i.e. they mainly care about the substance of decisions, and much less about the procedures leading to them (see also Arnesen 2017). These two positions nicely echo the tension, established by democracy theory, between input- and output-legitimacy (Scharpf 1999). Whereas 'input-oriented legitimacy' refers to the opportunities of individual citizens to participate in political decision making processes directly or indirectly through representatives, 'output-oriented legitimacy' highlights the substantive quality of decisions to "effectively promote the common welfare of the constituency in question" (Scharpf 1999, 6).

In this article, we explore citizens' preferences regarding democratic governance, focusing on their evaluations of a public policy according to criteria related to various legitimacy

dimensions, as well as on the (tense) relationship among them. To do so, we use experimental data from a conjoint analysis embedded in a cross-national survey conducted in France, Germany, Switzerland, and the UK. Our study thereby contributes to an emerging strand of political science research that uses experimental designs to examine the conditions of democratic legitimacy at the micro-level. This research, which we discuss in more detail in the next section, adopts an innovative approach to study the influence of decision-making arrangements on legitimacy beliefs. Our study enlarges this perspective in that we do not only focus on decision-making but on governance arrangements more broadly.<sup>1</sup>

The results suggest that citizens are predominantly concerned with output-legitimacy when evaluating governance arrangements. However, input- and throughput-legitimacy remain important secondary features. Interestingly, there is only scant evidence for citizens valuing interactions between the different legitimacy dimensions - thereby questioning a core argument of stealth democracy theory.

## DEMOCRATIC GOVERNANCE AND LEGITIMACY BELIEFS

We use the term ‘governance’ to designate the broad range of processes through which public policies are formulated and implemented. Governance thus involves performative goals against which it can be assessed by citizens, namely the production of policies that effectively tackle public problems. *Democratic* governance conveys the notion that the policy process conforms to a number of democratic norms. Both dimensions of democratic governance - the performative and the normative one - are thought to contribute to nurture legitimacy in the sense of citizens’ feelings that political power is rightfully held and exercised (Gilley 2006;

---

<sup>1</sup> With the term “governance arrangements” we refer to policy-making bodies vested with formal authority. These can be traditional state institutions and agencies, but also networks of organizations that are only loosely coupled to representative institutions. To use the terminology of Hooghe and Marks (2003), we refer to both type-I and type-II governance.

500). But what is their relative importance and how does their mutual relationship influence citizens' legitimacy beliefs?

In a first section, we start with a brief overview of existing research on the relative importance of democratic procedures compared to a system's performance for citizens' legitimacy beliefs. Based on this review, we posit that input, throughput and output matter for the generation of citizens' legitimacy beliefs in governance arrangements (see figure 1 for a graphical display of the hypotheses). In the second and the third section, we qualify this general assumption on two grounds. In line with an argument developed in stealth democracy theory (Hibbing and Theiss-Morse 2002), we posit that democratic processes are more important, when output evaluations are bad and become less important when output evaluations are good. Finally, drawing on arguments from accountability and international relations theories (Zürn et al. 2012, Warren 2014), we hypothesize that the higher a governance arrangement's formal authority, or more generally its "power", the more important democratic processes are for enhancing citizens' legitimacy beliefs.

[Figure 1 About Here]

Output, throughput and input: the basis for citizens' legitimacy beliefs

In political theory, it is a debated issue to what extent a political system's legitimacy depends on policy performance or on the conformity to democratic norms. Advocates of a substantive account in political theory argue that the normative value of democracy lies in the output it produces. Procedures alone are not sufficient to generate legitimacy and are mostly conceived as a means to an end (Arneson 2003). In this line of thought, democracy needs to "promote the common welfare of the people" (Bühlmann and Kriesi 2013, 44) to be legitimate.

Proceduralists, on the other hand, object that outcomes are always contested. For them, the

good quality of procedures is the only way for generating democratic legitimacy in an uncontested way (Dahl 1998).

Empirical political scientists are equally interested in the question of what renders a political system legitimate in the eyes of its constituency (Easton 1965; Norris 1999; Gilley 2009).

Two possibilities for political systems to enhance their subjective legitimacy (as perceived by constituencies) are traditionally conceived of: i.) allowing democratic input to the system or ii.) providing good output from the system. Drawing on Scharpf (1999: 6 ff.), we can define input-oriented legitimacy as an emphasis on ‘government by the people’, i.e. that political choices are assessed against the extent to which the individuals subject to these decisions were involved in their formulation either directly or indirectly through representatives.

Popular votes or direct election of representatives are examples of procedures that aim to ensure input legitimacy of political decisions.

By contrast, output-oriented legitimacy emphasizes ‘government for the people’, i.e. that political choices are assessed against the extent to which they effectively promote the common welfare of a community – the well-being of the people (Scharpf 1999). The effectiveness of policies in solving collective problems, or the efficiency of state agencies in delivering public services are examples of sources for output legitimacy. More recently, scholars introduced a third way for a political system to obtain legitimacy, namely “throughput”. Throughput legitimacy concerns the quality of (internal) governing processes of the institutions and actors concerned with policy-making (Schmidt 2013). Throughput-legitimacy can, for instance, be enhanced by a transparent way of policy-making – i.e. through public negotiation and decision-making processes (Héritier 2003) or by a good quality of deliberation in decision-making bodies – i.e. discussions and decisions based on reasons and not on political power considerations only (cf. Papadopoulos and Warin 2007). Throughput legitimacy thus “concentrates on what goes on inside the ‘black box’” between the input to, and the output from a political system (Schmidt 2013, 5).

Which of these three sources of legitimacy is driving citizens' evaluations of governance arrangements to the largest extent? Existing studies mainly distinguish procedural, i.e. input and throughput, and instrumental determinants, i.e. output, that drive citizens' legitimacy beliefs. For the evaluation of governance in the European Union, Fuchs (2011) finds that instrumental considerations (e.g. personal and collective benefits) are the most important correlates (cf. also Gabel and Palmer 1995). In contrast, Hooghe and Marks (2005) reveal that exclusive national identities – and associated concerns about self-determination and sovereignty – have a more negative effect on EU support than favorable economic calculus has a positive one. On the national level, an analysis of 54 third-wave democracies by Chu et al. (2008) show that citizens' support for the claim that democracy is the best form of government is not heavily affected by a bad evaluation of a regime's economic performance. At the same time, Dalton (2004, 74-76) finds that citizens' support for incumbents and for political institutions strongly depends on their assessment of their pocketbook. Rothstein (2009, 311) supports this finding and maintains that “[l]egitimacy turns out to be created, maintained, and destroyed not at the input but at the output side of the political system”. Moving to the local level, Denters (2014) finds that functional and procedural considerations are equally important for explaining satisfaction with local democracy. Yet, Perry (2014) demonstrates that satisfaction with the performance of local institutions is the most important driver of local attachment – an indicator for diffuse system support (cf. Norris 1999). From the international to the local level, citizens seem to be concerned with aspects of procedural and substantive quality when evaluating democratic systems.

While political scientists have a certain inclination to endorse the argument that system performance is the more important driver of legitimacy beliefs than democratic procedures (Hibbing and Theiss-Morse 2002, Esaiasson et al. 2016a, Arnesen 2017) – especially when it comes to specific support for political institutions and actors (cf. Easton 1965) – there is also evidence showing that opportunities for participation and active involvement in decision-

making as such are highly valued by citizens (Tyler 2000, Bengtsson and Mattila 2009, Esaiasson et al. 2012). It is thus difficult to formulate general expectations about which of these dimensions are more or less influential for legitimacy beliefs. The lowest common denominator – to which most political scientists can probably adhere – is to state that all three dimensions to a greater or lesser extent influence citizens' legitimacy beliefs in specific governance arrangements. We thus formulate three separate hypotheses:

H1a: Input considerations drive citizens' legitimacy beliefs in specific governance arrangements.

H1b: Throughput considerations drive citizens' legitimacy beliefs in specific governance arrangements.

H1c: Output considerations drive citizens' legitimacy beliefs in specific governance arrangements.

These hypotheses are very generic and they do not involve expectations about potential trade-offs between the different dimensions of democratic legitimacy. Yet, the stated aim of our paper is to shed light on the conditionality and the relative importance of different attributes of governance arrangements for citizens' legitimacy beliefs. In the next two subsections, we thus elaborate on the conditions under which democratic procedures are more or less relevant for citizens' legitimacy beliefs. On the one hand, we argue that bad output quality increases citizens' demand for democratic input and throughput, because they put greater emphasis on having an adequate possibility to voice their concerns and control policy-makers under these circumstances. On the other hand, we argue that the desire for input- and throughput-legitimacy is contingent on a governance arrangement's scope of authority: the more wide-ranging competences it possesses, the more citizens are affected by it and might want to co-determine its decisions.

## Output quality and the conditional importance of democratic procedures

The first qualification is simple: bad output. What happens if a governance arrangement does not meet the expected performance standards? To date, very few empirical studies have tackled this question (but see Lyons et al. 1992). However, a widespread argument that stems from stealth democracy theory goes like this: As long as political institutions deliver what is promised, citizens do not care about being involved in democratic processes. But as soon as performance is bad, citizens want to have the means to control policy-makers and to voice their claims. Hibbing and Theiss-Morse (2002, 105; 130) contend that Americans would rather have their fellow citizens making decisions than elected politicians when they believe politicians are not acting in their interest or try to personally benefit from their position. This can be interpreted as citizens' desire to have a corrective option in case things run off course – even if most of them might dislike political involvement in general. In an experimental study with Swedish high school students, Esaiasson et al. (2012) find that the possibility of direct majoritarian voting has a stronger effect on legitimacy beliefs than a fair implementation of the process. Even though their study design does not allow them to analyze interaction effects between the two conditions, their results could be interpreted in line with stealth democracy theory: In case things take a wrong turn, citizens want to have a channel to voice their claims. Neblo et al. (2010) add to the debate on stealth democracy and Hibbing and Theiss-Morse's (2002) basic claim according to which citizens do not want to be involved in political processes. In an experimental study on Americans' willingness to deliberate, they show that it is precisely individuals who are put off by “traditional politics” who are more willing to deliberate. Put differently, citizens dissatisfied with the way things work demand alternative channels of involvement in political processes. Finally, in two very recent experimental studies, Esaiasson et al. (2016a, 2016b) find that a.) outcome favorability – i.e. the correspondence between an individuals' preferences and the achieved outcome – influences

the perceived fairness of a procedure as well as the willingness to accept the outcome resulting from it and b.) outcomes that run counter to respondents' preferences are better accepted by them, when decision-makers take actions to inform themselves about citizens' preferences and to justify their decisions. Based on these rather scarce findings, we could expect that, when governance arrangements perform badly, citizens start to care (more) about possibilities to make their claims heard. Our second hypothesis thus reads as follows:

H2: When output quality worsens, the importance of input and throughput for citizens' legitimacy beliefs in specific governance arrangements increases.

### Formal authority and the conditional importance of democratic procedures

The second qualification to the first hypothesis concerns the scope of authority that a governance arrangement possesses. The relative importance of input, and throughput quality for citizens' evaluations of governance arrangements might not only depend on output quality, but also on the amount of decision-making authority a governance body possesses. This idea is based in an argument from the literature on democratic accountability: democratic influence on and control of decision-makers is particularly important when a governance body has wide-ranging competences (Warren 2014, 40). This argument is most prominent among scholars of international politics (cf. Mulgan 2014, 546). In this view, international organizations have a democratic deficit, because input possibilities for citizens are reduced and representation becomes more indirect as formal authority moves away from the representative institutions of the nation state (Dahl 1999). However, this democracy deficit is not only present in international politics. The same is true when public service provision is outsourced or regulation authority is shifted from national government agencies to private contractors or to independent regulatory agencies (Mulgan 2014, 549). Such developments also entail a strengthening of formal authority of these bodies paralleled by a weakening of

democratic accountability. This in turn might trigger demands for more democratic input and throughput in decision-making processes.

Again, empirical research on citizens' perceptions of these issues is scarce. In fact, we have only scant and very indirect evidence that an asymmetrical expansion of formal authority without a simultaneous strengthening of democratic procedures renders mass publics more critical of such institutions. In a study on the politicization of international organizations, Zürn et al. (2012, 96) argue that the more authority an international organization possesses, the more contested and politicized are its actions and the more debated is its democratic legitimacy. Moreover, studies on public support for European integration suggest that a negative relationship between levels of formal authority and legitimacy perceptions might indeed exist. Citizens generally have become more critical of European integration in the post-Maastricht period, the time since when the scope and depth of EU decision-making authority expanded considerably (e.g. Hooghe and Marks 2008, 10). In short, as the formal authority of international organizations increases democratic decision-making processes and input opportunities become more important, also from a citizens' perspective.

This mechanism is not limited to international organizations. We argue that an increase in the formal authority of any governance body goes along with a heightened demand for democratic input possibilities. When a governance arrangement obtains more decision-making authority, citizens might also want to be more involved in the respective decision-making processes. This leads us to our last hypothesis:

H3: When formal authority increases, the importance of input and throughput for citizens' legitimacy beliefs in specific governance arrangements increases.

## RESEARCH DESIGN

We use an experimental design to test these hypotheses. Since we are interested in the relationship between different attributes of governance arrangements that drive citizens' evaluations, an experimental design that allows a random variation of these attributes is the best strategy for studying how these variations affect citizens' legitimacy beliefs. Moreover, experiments allow to overcome the endogeneity problems of studies that use natural cross-institutional variations or observational data to investigate citizens' preferences. Our study thereby relates to an emerging strand of literature in political science that uses experimental approaches to investigate citizens' legitimacy beliefs and their assessment of various decision-making arrangements (see Tyler 1994, Esaiasson et al. 2012, Arnesen 2017).

However, our study is specific in two respects. On the one hand, we do not use field experiments (as in Tyler 1994 or Esaiasson et al. 2012, 2016b) but an experiment embedded in a population-based survey.<sup>2</sup> A limitation of traditional experimental designs is that they only allow for assessing the effect of one treatment at a time. However, our endeavor is to analyze citizens' reaction to different attributes that vary simultaneously. Put differently, the treatment consists of multiple components which all are expected to have an impact. To analyze the effects of such multi-dimensional treatments, Hainmueller et al. (2014, 2) propose conjoint analysis "as a tool to identify the causal effects of various components of a treatment in survey experiments." In our case, the attributes of a governance arrangement are the components, whose effects on citizens' evaluations we want to assess and conjoint analysis is thus a well-suited experimental design for this endeavor.

Another specificity of our study is its scope. Previous experimental studies on similar topics focus on how various modes of decision-making in small scale settings affect individual

---

<sup>2</sup> The only other study we have identified that tests the importance of procedures and outcomes for citizens decision acceptance in a population-based survey stems from Esaiasson et al. (2016a). Yet, their survey experiment is limited to the context of metropolitan Gothenburg.

assessments of procedural fairness (Tyler 1994, Esaiasson et al. 2016b), or decision-acceptability (Esaiasson et al. 2016a, Arnesen 2017). Given this narrow focus on decision-making, it is unclear to what extent the findings from these studies are valid for the larger scale of political systems more generally. This is why our study focuses more broadly on governance arrangements, which involve not only decision-making, but also the implementation of these decisions.

The context: metropolitan governance arrangements in public transportation

This subsection discusses the context and the precise features of the governance arrangement assessed here. We focus our analysis on governance arrangements in metropolitan areas, i.e. vast functionally integrated urban settlements, also called city-regions.<sup>3</sup> Metropolitan areas are not only the core sites of the 21<sup>st</sup> century capitalist economy (Brenner 2003), they are also the main habitat of the human species since 2010 (United Nations 2014). The ongoing expansion of metropolitan areas worldwide creates challenges for their governance. An increasing number of issues require solutions that transcend local boundaries and require the cooperation of governments from various territorial levels (Heinelt and Kübler 2005; Kübler and Pagano 2012). In other words: metropolitan areas are particularly meaningful places to study the workings of governance arrangements that seek to address major policy issues.

We center our analysis on the issue of public transportation. Public transportation is a paradigmatic case of a metropolitan problem that requires coordination between different political actors in metropolitan areas (Gerber and Gibson 2009, 635). The main purpose of public transportation is to connect places and enable a quick movement of people from A to B. This increased connectivity in turn results in higher degrees of functional territorial

---

<sup>3</sup> As a rule, statistical offices define metropolitan areas on the basis of commuting patterns between the cities, towns and villages in an area (Hoffmann-Martinot and Sellers 2005). However, the precise definitions of metropolitan areas vary across different countries. For having a comparable definition of metropolitan areas, we therefore rely on Eurostat's (2013) "functional urban areas" definition.

integration because it allows people to live in one place and work in another. The performance of the public transportation system is thus pivotal for the well-being of metropolitan populations.

What can this rather specific case tell us about our general hypotheses? Governance at the local and the metropolitan level is mostly about “getting things done” in an efficient way (cf. Haus 2014). Local politics are said to deal with idiosyncratic issues beyond partisan cleavages and political ideologies only play a minor role compared to the national level (Oliver 2012). If local governance is depoliticized, this is even more the case for metropolitan governance (cf. Deas 2014). As a rule, citizens are not very familiar with metropolitan governance arrangements and their knowledge about political processes at this level is limited (Swianiewicz and Lackowska 2007). If citizens show little interest in being involved in politics at the national level (Hibbing and Theiss-Morse 2002), this might even more be the case for politics at the local and the metropolitan level. This means that we are analyzing governance arrangements in a context in which citizens have an elevated disposition to choose output over input.

A similar assessment can be made for the issue of public transportation. In contrast to metropolitan governance in general, citizens certainly are familiar with the policy area of public transportation. The quality of public transportation affects the daily life of many metropolitan residents and they thus have a vital stake in how public transport services work. However, unlike policies involving redistributive questions that are conducive to ideological considerations – which renders procedural questions more relevant – public transport governance is biased towards providing a service in an efficient way. It is thus plausible to assume that citizens do not necessarily care about whether they can influence policy decisions

as long as public transportation works at a reasonable cost. Like local and metropolitan governance in general, public transportation in particular is about “getting things done”.<sup>4</sup> What does this mean for the generalizability of our findings? It would, obviously, not be a surprise if we find that output matters more than input or throughput for citizens’ evaluation of public transport governance in metropolitan areas. In this sense, we are analyzing a most likely case for output to matter, and for input and throughput to be less relevant (cf. Gerring 2007). By contrast, if we find that input and throughput matter nonetheless, this would have rather strong implications for governance in other policy fields and on other levels as well.

The scenario: a commission for a new underground line in the metropolis

The experimental design we chose was to confront respondents with a variety of governance set-ups aimed at improving the public transport network in their metropolitan area, and ask them to indicate which one of those set-ups they preferred. As outlined in table 1, this was done in two steps. In a first step, respondents were told to imagine that public authorities in their metropolitan area decided that a new underground train line shall be built in their region in order to solve traffic problems, and that a commission<sup>5</sup> is installed to define, design and carry out this project.

In a second step, respondents’ preferences for commissions varying on five attributes (see table 1) were measured. More precisely, each respondent was presented with three choice tasks of opposing pairs of different commissions, in which she had to indicate the one of the two commissions she preferred. In addition, respondents were asked to rate each commission

---

<sup>4</sup> Of course, one can think of situations in which public transportation policies involve issues of redistribution and ideological conflict. For example, debates about (the level of) subsidies for public transport tickets by the local or regional government both have a redistributive and an ideological component: There is redistribution from the average tax payer to the public transport user, which can provoke debates about state interventionism and public finance with a clear ideological underpinning. In the scenario that we presented to respondents, there was however a clear bias towards “getting things done” and towards efficient governing. Redistributive and ideological issues were not made salient to the respondents in our scenario (see below).

<sup>5</sup> The term ‘commission’ was deliberately vague, to avoid that respondents make references to existing agencies or authorities in their metropolitan area of residence.

on a scale from 1-7 (see table 2 for an example).<sup>6</sup> Each commission consisted of a random combination of attribute levels and the order of the five attributes was randomized across respondents but remained constant within respondents.<sup>7</sup>

While the choice and the rating of a commission represent the dependent variables, the five attributes constitute the independent variables in the conjoint analysis. Three of these attributes operationalize different dimensions of democratic legitimacy (specified in hypotheses 1 and 2) and two of them operationalize the formal authority a commission holds (specified in hypothesis 3).

The first attribute refers to the composition of a commission and varies the degree of its input-legitimacy. This attribute has three levels: decision-makers could be directly elected by citizens, delegated by local authorities, or appointed independent technical experts. Elected members represent the most direct possibility for input by the citizens, and thus reflect the idea of a democratically elected metropolitan government in scholarly debates (Lefèvre 1998; Gerber and Gibson 2009). Delegate members are only indirectly linked to the citizens. This composition reflects inter-municipal cooperation arrangements in many European countries (Hulst and Van Montfort 2007). Finally, an expert commission is a technocratic body, not accountable to citizens and thus has the least input legitimacy.<sup>8</sup> The second attribute aims to

---

<sup>6</sup> In a validation of different types of conjoint and vignette experiments against real-world behavior, Hainmueller et al. (2015) demonstrate that this variant – a paired conjoint analysis – most closely mirrors citizens’ decision-making in real-world situations.

<sup>7</sup> This can prevent “attribute order effects” – effects that occur only because of the position of an attribute.

<sup>8</sup> As one of the anonymous reviewers rightly pointed out, respondents will assume that the independent experts are appointed by a democratically elected authority, the “public authorities” from our scenario who install the commission in the first place. Therefore, respondents might not perceive these independent experts to mean less involvement on their part than the delegates of the local authorities. In the specific case at hand, there’s still an important difference between the two, however: In all the metropolitan areas under scrutiny here (see below), public authorities with the competence to decide on such a large-scale public transport project are located at the supra-local scale. Local governments are not directly involved in such a decision. An important difference between the local delegates and the independent experts is thus whether they are appointed bottom-up, by the municipalities, or top-down, by the higher government tier. There is thus less *local* involvement in the latter case. While we can only assume that respondents know which public authorities are responsible for making such a decision, this multilevel aspect of metropolitan politics is an often salient dimension of conflict in metropolitan areas (Gerber and Gibson 2009, Kübler and Pagano 2012). An implication of this ‘two-step’ operationalization of democratic involvement, (democratically elected) public authorities (step 1) install a commission (step 2) is that the input dimension, i.e. the composition of the commission, might be regarded as even less relevant by respondents, since there’s already the possibility for democratic involvement at step 1. For our findings, this

capture throughput legitimacy. A central criterion for throughput legitimacy is the transparency of the negotiation- and the decision-making process (Héritier 2003). Only this allows citizens to evaluate the different arguments made and hold the respective actors accountable for their actions (Haus 2014). Therefore, we operationalize this dimension by two attribute levels that indicate whether these decision-making processes were public or not. Finally, the third attribute - the degree of budget-overrun - captures the efficiency of a commission's work and thus one important dimension of output legitimacy (cf. Bernauer et al. 2016). While one might argue that the effectiveness – the actual functioning of the new underground line – could be a more important aspect of output legitimacy, we decided to use cost-efficiency as a proxy for output legitimacy for three reasons. First, operationalizing the effectiveness of the commission's output as to whether the new underground line works or not, would have been too strong of a treatment. When the sole purpose of the commission is to design, plan and build a new underground line and it completely fails to do so, the very *raison d'être* of the commission is undermined. Respondents would consequently only evaluate this attribute and ignore the others. This is why the story introducing our experiment was formulated in a way that respondents had to assume that actual completion of the metro line was not at stake, and that it did serve its purpose i.e. to alleviate transport problems in the respondent's metropolitan area. Second, problems with budget over-runs are a common feature of large infrastructure projects and, if they occur, are usually widely debated in public – and often characterized as a bummer. An operationalization of the output-dimension via the cost-efficiency of the infrastructure project thus closely resembles real-world situations and gives respondents a clear signal that is familiar to them. Finally, in studies on the performance and output legitimacy of local governance arrangements cost-efficiency figures as a prominent way of operationalizing output quality (Pierre 2011) and some local governance scholars even equate output legitimacy with efficiency (Kersting & Vetter 2003). For these

---

means that, if we still find an impact of a commission's composition on respondents' choice, this would be a strong indication that input possibilities matter in other contexts as well.

reasons, we varied the extent to which the projected budget was over-run with three degrees: 0 to 10 percent, 10 to 20 percent, 20 to 30 percent.

[Table 1 About Here]

For the operationalization of formal authority, we relied on two indicators derived from Hooghe and Marks' (2015, 315) measurement of delegation and pooling of formal authority in multi-levelled organizations. Both of our attributes belong to the "pooling"-dimension. This is the case because "[d]elegation is a grant of authority to a third party" (Hooghe & Marks 2015, 315) Delegation thus refers to the tasks of an organization. This is already captured in our introductory story, i.e. public authorities installing a commission with a precise set of tasks, and is not part of the randomized treatment. Hooghe and Marks (2015) operationalize pooling with three indicators that concern the way in which final decisions are taken in a multi-levelled organization: decision rule (unanimity vs. majority), need for ratification and bindingness. We decided to combine the "bindingness" and the "ratification" of a decision into one single attribute which captures the veto power of local governments. We have done so first to reduce the complexity of the experiment<sup>9</sup> and second, to prevent non-logical combinations of attribute levels<sup>10</sup>.

[Table 2 About Here]

These five attributes and their two- to three attribute levels yield a total of 72 possible commissions ( $2^3 * 3^2 = 72$ ) and 2556 possible combinations of commissions ( $72 * 71/2 = 2556$ ). A total of 5052 respondents participated in our survey (see below) which means that – on

---

<sup>9</sup> For example, the difference between a decision not being binding for the involved actors and a decision having to be ratified by them is a very subtle one that is difficult to grasp for respondents.

<sup>10</sup> For example, it could be perceived as a contradiction for a decision to be binding for and to require ratification by local authorities at the same time.

average – each possible combination was evaluated almost 12 times ( $5052 \cdot 6 / 2556 = 11.8$ ) and each commission was on average evaluated 421 times ( $5052 \cdot 6 / 72 = 421$ ). In contrast to other conjoint analyses – in which only a sample of possible combinations is assessed (Hainmueller et al. 2014) – each combination was assessed several times in our case. This increases the internal validity of our results.

The data: a representative population survey in eight European metropolitan areas

The data for this study stem from an online-survey that was conducted in eight metropolitan areas in four countries. It was fielded in fall 2015 and includes a total of 5052 interviews of representative samples of the resident population aged 18 to 75 in two Swiss, two German, two French and two British metropolitan areas (for details on the sampling strategy and recruitment of respondents, see appendix A in online supplementary materials).

[Table 3 About Here]

The distribution of respondents within metropolitan areas reflects the spatial distribution between center city and suburban areas among the general population of these areas (see table 3). The countries and the metropolitan areas to a certain extent reflect a most-different case selection (Gerring 2007). The countries differ with respect to traditions of local and metropolitan governance (Heinelt and Kübler 2005), as well as the authority and autonomy of subnational government tiers (Ladner et al. 2015, Hooghe et al. 2016). In each country the capital region and one other major metropolitan area was chosen. Finally, the metropolitan areas also differ in terms of their population size. This means, that if we find similar results across these different cases, this provides a strong basis for a generalization of the findings beyond the eight metropolitan areas and to other European countries.

## Estimation strategy

For our analysis we rely on the `-cjoint-` package in R developed by Hainmueller et al. (2014). As reported above, each respondent was presented with three choice tasks and accordingly evaluated six governance arrangements, (i.e. “commissions”). This means that choices are nested within respondents requiring a transformation of the dataset: For each respondent we do not have one but six rows in the dataset - one for each commission that she evaluated. The unit of analysis is thus one respondent-commission. To account for the nesting of commissions within respondents, we use clustered standard errors (Hainmueller et al. 2014). The independent variables – i.e. the five attributes – are introduced to the regression equation as factor variables. This means that one of the attribute levels of each attribute serves as the baseline. We, furthermore, use poststratification weights based on age, sex, education, employment status and place of residence within the metropolitan area to correct for survey non-response bias (see Table A.1).

The estimates obtained through the use of the `-cjoint-` package are *average marginal component effects* (Hainmueller et al. 2014). These estimates can be interpreted as the effect of an individual treatment *component* – the whole commission or attribute combination being the treatment. More precisely, “the *average marginal component effect* (AMCE) represents the marginal effect of attribute *l* averaged over the joint distribution of the remaining attributes” (Hainmueller et al. 2014, 10). For the first dependent variable – the forced choice of one of the two commissions – this quantity can be interpreted as the percentage point change in the probability of choosing a commission when a certain attribute changes from the baseline level to the level of interest. For the second dependent variable – the rating of a commission – this quantity represents the change on the 1-7 scale when an attribute changes from the baseline level to the level of interest. In what follows, we will interpret the effects of

the first dependent variable. The substantive results for the rating variable do not differ from the results for the forced choice variable (see appendix E, supplementary materials). For the analysis, we deleted rows with missing values for the dependent variables listwise. The number of missing values is very low compared to the overall number of rows (114 rows out of 30312) and should thus not bias the results.<sup>11</sup>

We subjected the findings to a number of robustness checks and tested whether carryover effects occur (appendix B), whether results vary across respondent characteristics (appendix C), across response types (appendix D) and across countries and metropolitan areas (appendix F). Our results remain stable across all of these alternative specifications.

---

<sup>11</sup> This corresponds to a reduction in the number of respondents from 5052 to 5040. Twelve respondents did not complete a single choice task. The remaining 42 missing rows are due to respondents not completing one or two of the three choice tasks. We decided to keep them in the sample, since the completed choice tasks still generate valid data.

## RESULTS

In this section we present the main results of our conjoint experiment and test the three hypotheses: i.) that the input- (H1a), the throughput- (H1b), and the output-dimension (H1c) matter for the evaluation of specific governance arrangements by citizens, ii.) that citizens' evaluation of input and throughput aspects is contingent on output quality and iii.) that a commission's level of formal authority conditions citizens' evaluation of input and throughput aspects.

Hypotheses 1a, b, and c: input, throughput, and output matter

Figure 2 presents the results for the basic analysis of our conjoint experiment, which allows us to test our first set of hypotheses. The results show that in their assessment of governance commissions seeking to improve public transportation citizens prefer more democratic input (elected > delegated commission members), they prefer more democratic throughput (public > non-public negotiations) and they prefer better output (0-10% > 10-20% > 20-30% budget over-run). Interestingly, commissions composed of delegates are even less preferred than commissions composed of experts. This goes against our expectation, which states that delegates mean more democratic involvement than experts. Respondents do not seem to respond to the, admittedly rather subtle, difference between delegates appointed bottom-up by local governments and experts appointed top-down by the public authorities in the region. Moreover, since the task of the commission mostly lies in implementing a public policy decision, respondents might also acknowledge that experts would be the persons more qualified for this task than local delegates.

In sum, we find support for our first set of hypotheses. Democratic input (H1a) and throughput (H1b), as well as cost-efficient output (H1c) matter for citizens' legitimacy beliefs. Moreover, however, cost-efficient output seems to matter much more for respondents' choices than more democratic input (the 0-10% budget overrun category indeed shows the highest AMCE of all attribute levels). Throughput – in the form of transparency – figures somewhere in between. A commission with only 0-10% budget over-run is more than 17 percentage points more likely to be chosen by respondents than a commission whose work results in a 20-30% budget over-run. At the same time, a commission composed of directly elected members is only 2.5 percentage points more likely to be chosen than a commission composed of experts.

[Figure 2 About Here]

Interestingly, the throughput-dimension seems to be the second most important one to determine respondents' choice of a commission, whereas the mode of choosing decision-makers for a commission (election, delegation or appointment) is considered to be the least important attribute, even less important than the two formal authority dimensions. Finally, commissions that were shown on the left side of the screen are preferred over commissions shown on the right side. We thus need to include the control variable for these “profile-order-effects” (Hainmueller et al. 2014, 9) in subsequent analyses as well.

Hypotheses 2: a conditioning effect of output quality on democratic input and throughput?

So far, we have found support for our baseline hypothesis, that democratic input, throughput and efficient output are influencing respondents' preferences for governance arrangements.

We now turn to the test of hypotheses 2, namely that the relative importance of input and throughput increases when a commission's work results in a higher budget over-run.

To do so, we interact the respective attributes with one another. The results for these interactions are shown in figure 3. The relative importance of democratic input and throughput does not increase significantly, as budget over-run increases. While the effect of having transparent negotiations remains constant across the three levels of budget over-run, it seems that a commission composed of experts is evaluated even a bit more favorably, when budget over-run is high (20-30%) compared to when it's low. These changes in the effect size across different levels of budget over-run are almost significantly different from zero with 95% confidence. Hypothesis 2 therefore needs to be rejected: The preference of citizens for commissions that emphasize democratic input and throughput does not increase when a commission's performance is bad. To the contrary, it seems that citizens prefer what can be seen as more efficient input, i.e. appointing experts rather than electing representatives, when a commission's performance is bad. This might reflect a belief that democratic input is a "luxury" good one is willing to abandon for the sake of more efficient output, especially when the main issue at stake is the implementation of a policy.

[Figure 3 About Here]

Hypothesis 3: a conditioning effect of formal authority on democratic input and throughput?

The third hypothesis states that an increase in formal authority also increases the relative importance of democratic input and throughput in citizens' assessment of governance arrangements. Figures 4 and 5 contain the interaction effect between our formal authority measures, decision mode and implementation, and our input and throughput measures.

Judging from figure 4, there does not seem to be any interaction effect between higher formal authority of a commission (=majoritarian decision-making) and the relative importance of democratic input and throughput. While a commission composed of directly elected members

seems to be slightly more preferred over experts when the commission operates with a majoritarian rather than a unanimous decision mode, this difference is far from significant. The importance of public compared to non-public negotiations remains constant across decision modes.

[Figure 4 About Here]

Finally, figure 5 displays the interactions between the second attribute that captures the formal authority of a commission, namely whether the final decision requires approval or not, and input and throughput. In contrast to the first attribute of formal authority, we find partial evidence for our third hypothesis here. Figure 5 shows that the positive effect of having directly elected representatives instead of experts decreases by 2.5 percentage points when commissions' decisions require approval compared to when they don't require approval – effectively rendering the difference between elected representatives and experts insignificant. This means that the extent to which the final decision of a governance body requires approval by the affected constituencies seems to condition the importance of democratic input. However, the same is not true for more democratic throughput. The desire for public negotiations does not covary with the extent of a commission's implementation authority.

[Figure 5 About Here]

## Discussion

In sum, we find support for our first set of hypotheses. More democratic input and throughput as well as efficient output increase respondents' acceptance of a commission. Furthermore, the cost-efficiency (the proxy for output) of the final product – i.e. the constructed

underground line – is a more important driver of respondents’ choices than the composition of the commission (input) or the way it communicates with the public (throughput). This is not a very surprising finding, given the fact that both the level at which decisions are made (metropolitan) as well as the issue area (public transportation) and the present scenario have a clear penchant for “getting things done”. Political ideology and questions of redistribution only play a minor role. What is more surprising, however, is that we still find significant positive effects of more democratic input and throughput on the choice of a commission under these conditions – even when a commission’s performance is very good. On the basis of this finding, we have to reject our second hypothesis. Moreover, this finding has implications for the debate on stealth democracy: citizens seem to want democratic input – and even slightly more so when things are running well. This runs against a central tenet of stealth democratic reasoning: that most citizens do not care about being involved in democratic procedures as long as they get the outcomes they want (Hibbing and Theiss-Morse 2002, 143). This finding is in line with other recent studies on the relation between input and output legitimacy. In an experimental study among Norwegian citizens, Arnesen (2017) finds no interaction between respondents’ possibility to influence an outcome and whether or not a respondent’s outcome preferences were met by the decision on their propensity to accept the decision. Similarly, in a conjoint experiment on citizens’ preferences for global environmental governance, Bernauer et al. (2016) find no interaction between output quality and input possibilities either.

The findings for the third hypothesis are mixed. There’s no evidence of a conditioning effect of the way decisions are made within a commission on the importance of democratic input and throughput, which contradicts the idea that higher formal authority of a governance arrangement increases demand for democratic input and throughput. Yet, when it comes to the implementation of a commission’s decision, matters are different. Here, democratic input becomes more important when a commission has wide-ranging formal authority to implement its decisions. The fact that we find such an interaction effect in a largely depoliticized setting,

both in terms of context and policy area, suggests that such an interaction might also be present – and probably stronger – in more politicized contexts.

## CONCLUSION

On what grounds do citizens' evaluate the legitimacy of specific governance arrangements?

Our results suggest that citizens' generally value governance bodies that would be considered democratic from a "traditional" representative understanding: They demand democratic input and throughput as well as cost-efficient output and they favor governance arrangements whose internal functioning works according to a majoritarian logic but whose implementation authority is restricted by external actors. Cost-efficiency is the most important driver of citizens' evaluations of metropolitan governance arrangements among the dimensions tested in our experiment. To a certain extent, this is in line with Scharpf's (1999) famous claim that problem-solving capacity is the most important tool for political actors to create democratic legitimacy perceptions among their constituents. Contrary to our expectations, there is no evidence for an interaction effect between output efficiency and democratic input and throughput. Even in a setting that is highly geared towards solving public problems and even when these problems are solved efficiently, democratic input and especially throughput remain important drivers of citizens' evaluations of governance arrangements. This finding contradicts a central tenet of stealth democratic theory, that, given desirable outputs, citizens do not care about being involved in democratic processes (Hibbing and Theiss-Morse 2002, Arnesen 2017). Instead, we find that citizens do care about the democratic quality of decision-making procedures, no matter if the outputs are desirable or not. In this sense, our study confirms earlier findings from experiments in small-scale settings, according to which individuals value the substantive output of decisions, but simultaneously have intrinsic preferences for procedures that allow their participation in decision-making - independently of

whether their participation is instrumental to shaping this output (Tyler 1994, 2000, Esaiasson et al. 2012).

Our study has important limitations, however. First, the result that output drives citizens' evaluations of governance arrangements more strongly than input or throughput is not astonishing, given the fact that we are dealing with a scenario which is mostly about solving a problem in an efficient way. Results might be different on other governance levels and in other policy areas, where one can expect a stronger effect of democratic input and throughput on citizens' evaluations. Second, however, our output treatment is rather "tame". While cost-efficiency is an important aspect of public policy outputs – especially in today's context of fiscal austerity – the effectiveness of a policy in solving public problems is at least as important. It is conceivable that a manipulation of that aspect of output might have brought about more significant interaction effects of output with democratic input and throughput. Finally, our experimental scenario is about implementing a public policy that has been decided on in an earlier step. This might explain why we find that experts are preferred as commission members over local delegates, even if that means less democratic involvement of citizens. The involvement has already happened at an earlier stage and is considered less important for implementation. We might thus expect that in another scenario that is more about policy-making rather than implementation, respondents might be more sensitive to such differences and they might also demand more democratic involvement when their preferences do not match the outcome (cf. Esaiasson 2016a).

Future research should therefore scrutinize our results for different policy areas and policy types in different contexts. Particularly promising would be experiments concerning welfare state reform. While some scholars engaged in conjoint experiments in this domain (see Gallego and Marx 2016), they are concerned with the design of a policy and not with the evaluation of the institution that is responsible for developing it. In addition to the scope conditions of our findings, the links between input, throughput and output legitimacy have to

be explored further from a citizens' perspective. While our results suggest that these three central aspects of democratic legitimacy operate largely independently from one another in subjective legitimacy evaluations, other studies suggest that citizens might also view synergies between them (Gustavsen et al. 2014). The design of our study does not allow for testing whether citizens believe that positive interaction effects exist between democratic input, throughput and output, e.g. that more democratic input also produces better output. This is something future research should explore in more detail.

Finally, our results also have practical implications. While they suggest that from a citizens' perspective the most important quality of political institutions is the capacity of "getting things done" and that procedural traits of liberal democracy are only of secondary importance, they also show that procedural traits are *not* irrelevant. To obtain political support, it is not enough to have a technocratic government that provides efficient output and ignores the procedural and pluralist side of democratic politics (cf. Caramani 2017) – not even in the case of a pareto-efficient public policy problem like underground line construction in metropolitan areas. The positive interaction between a commission's leeway to implement its decisions and respondents' valuing of democratic input possibilities suggests that more task-specific, technocratic expert governance – i.e. type II – is tolerated only within a framework which still allows for democratic control – and thus bears type I features (cf. Hooghe and Marks 2003). While procedures are not (always) citizens' prime concern, they remain important features of legitimate governance arrangements. This is a promising avenue for contemporary democracies in general and indicates that we should strive to improve not only the quality of public services but also the quality of democratic procedures.

## ACKNOWLEDGEMENTS

Research for this paper was conducted in the framework of the National Centre of Competence in Research “Challenges to Democracy in the 21<sup>st</sup> Century” at the University of Zurich and funded by the Swiss National Science Foundation. Previous versions of this paper were presented at the 10<sup>th</sup> ECPR General Conference in Prague 2016, at the joint meeting of the German, the Austrian and the Swiss Political Science Associations in Heidelberg 2016, at the 9<sup>th</sup> NCCR Democracy Internal Conference in Thun 2016, at the Final Conference of the NCCR Democracy in Zurich 2017, as well as at the Séminaire général du Centre d'études européennes de Sciences Po Paris in 2017. The authors would like to thank Anja Durovic, Joe Greenwood, Lea Heyne, Gary Marks, Nicolas Sauger, Markus Tausendpfund, and Werner Wirth, the participants of these conferences, as well as the editors and the two anonymous referees of the European Journal of Political Research for their most valuable comments and suggestions.

## REFERENCES

Arnesen, S. (2017). Legitimacy from decision-making influence and outcome favourability: results from general population survey experiments. *Political Studies* 65(1S): 146-161.

Arneson, R.J. (2003). Debate: defending the purely instrumental account of democratic legitimacy. *The Journal of Political Philosophy* 11(1): 122-132.

Bengtsson, A. & Mattila, M. (2009) Direct democracy and its critics: support for direct democracy and 'stealth' democracy in Finland. *West European Politics* 32(5): 1031-1048.

Bernauer, T., Koubi, V. & Mohrenberg, S. (2016). How relevant are input and output legitimacy in international environmental governance? *NCCR Working Paper* 94. Available at: [http://www.nccr-democracy.uzh.ch/publications/workingpaper/pdf/wp\\_94.pdf](http://www.nccr-democracy.uzh.ch/publications/workingpaper/pdf/wp_94.pdf) [Accessed: 29 November 2017].

Brenner, N. (2003). Metropolitan institutional reform and the rescaling of state space in contemporary Western Europe. *European Urban and Regional Studies* 10(4): 297-324.

Bühlmann, M. & Kriesi, H. (2013). Models for democracy. In H. Kriesi et al. (eds.), *Democracy in the age of globalization and mediatization*. Basingstoke: Palgrave Macmillan, pp. 44-68.

Caramani, D. (2017). Will vs. reason: the populist and technocratic forms of representation and their critique to party government. *American Political Science Review* 111(1): 54-67.

Chu, Y.H. et al. (2008). Public opinion and democratic legitimacy. *Journal of Democracy* 19(2): 74-87.

Dahl, R.A. (1998). *On democracy*. New Haven: Yale University Press.

Dahl R.A. (1999). Can international organizations be democratic? A skeptic's view. In I. Shapiro & C. Hacker-Cordón (eds.), *Democracy's edges*. Cambridge: Cambridge University Press, pp. 19-37.

Dalton, R.J. (2004). *Democratic challenges, democratic choices. The erosion of political support in advanced industrial democracies*. Oxford: Oxford University Press.

Deas, I. (2014). The search for territorial fixes in subnational governance: city-regions and the disputed emergence of post-political consensus in Manchester, England. *Urban Studies* 51(11): 2285-2314.

Denters, B. (2014). Beyond 'what do I get?' Functional and procedural sources of Dutch citizens' satisfaction with local democracy. *Urban Research & Practice* 7(2): 153-168.

Easton, D. (1965). *A systems analysis of political life*. New York: Wiley.

Esaiasson, P., Gilljam, M. & Persson, M. (2012). Which decision-making arrangements generate the strongest legitimacy beliefs? Evidence from a randomised field experiment. *European Journal of Political Research* 51: 785-808.

Esaiasson, P., Gilljam, M., Persson, M. & Lindholm, T. (2016a). Reconsidering the Role of Procedures for Decision Acceptance. *British Journal of Political Science*, doi: 10.1017/S0007123416000508.

Esaiasson, P., Gilljam, M. & Persson, M. (2016b). Responsiveness beyond policy satisfaction: does it matter to citizens? *Comparative Political Studies* 50(6): 739-765.

Eurostat (2013). *Territorial typologies for European cities and metropolitan regions*.

Available at: [http://ec.europa.eu/eurostat/statistics-explained/index.php/Territorial typologies for European cities and metropolitan regions](http://ec.europa.eu/eurostat/statistics-explained/index.php/Territorial_typologies_for_European_cities_and_metropolitan_regions)

[Accessed 29 November 2017].

Fuchs, D. (2011). Explaining support for European integration: an attitudinal model. In D. Fuchs & H.D. Klingemann (eds.), *Cultural diversity, European identity and the legitimacy of the EU*. Cheltenham: Edward Elgar, pp. 220-246.

Gabel, M. & Palmer, H.D. (1995). Understanding variation in public support for European integration. *European Journal of Political Research* 27: 3-19.

Gallego A. & Marx, P. 2017. Multi-dimensional preferences for labour market reforms: a conjoint experiment. *Journal of European Public Policy* 24(7): 1027-1047.

Gerber, E.R. & Gibson, C.C. (2009). Balancing regionalism and localism: how institutions and incentives shape American transportation policy. *American Journal of Political Science* 53(3): 633-648.

Gerring, J. (2007). *Case study research. Principles and practices*. Cambridge: Cambridge University Press.

Gilley, B. 2006. The meaning and measure of state legitimacy: results for 72 countries. *European Journal of Political Research* 45: 499-525.

Gilley, B. (2009). *The right to rule. How states win and lose legitimacy*. New York: Columbia University Press.

Gustavsen, A., Røiseland, A. & Pierre, J. (2014). Procedure or performance? Assessing citizen's attitudes toward legitimacy in Swedish and Norwegian local government. *Urban Research & Practice* 7(2): 200-212.

Hainmueller, J., Yamamoto, T. & Hopkins, D.J. (2014). Causal inference in conjoint analysis: understanding multidimensional choices via stated preference experiments. *Political Analysis* 22: 1-30.

Hainmueller, J., Hangartner, D. & Yamamoto, T. (2015). Validating vignette and conjoint analysis against real-world behavior. *Proceedings of the National Academy of Sciences* 112(8): 2395-2400.

Haus, M. (2014). Mirror of the state or independent image? Conceptual perspectives on the question of a legitimacy shift to the output dimension in local democracy. *Urban Research & Practice* 7(2): 123-136.

Heinelt, H., & Kübler, D. (2005). *Metropolitan governance. Capacity, democracy and the dynamics of place*. London: Routledge.

Héritier, A. (2003). Composite democracy in Europe: the role of transparency and access to information. *Journal of European Public Policy* 10(5): 814-833.

Hibbing, J.R. & Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. Cambridge: Cambridge University Press.

Hoffmann-Martinot, V. & Sellers, J. (Eds.) (2005) *Metropolitanization and political change*. Opladen, Verlag für Sozialwissenschaften.

Hooghe, L. & Marks, G. (2003). Unraveling the central state, but how? Types of multi-level governance. *American Political Science Review* 97(2): 233-243.

Hooghe, L. & Marks, G. (2005). Calculation, community and cues. *European Union Politics* 6(4): 419-443.

Hooghe, L. & Marks, G. (2008). A postfunctionalist theory of European integration: From permissive consensus to constraining dissensus. *British Journal of Political Science* 39(1): 1-23.

Hooghe, L. & Marks, G. (2015). Delegation and pooling in international organizations. *Review of International Organizations* 10(3): 305-328.

Hooghe, L., et al. (2016). *Measuring regional authority: A postfunctionalist theory of governance, volume I*. Oxford: Oxford University Press.

Hulst, R. & Van Montfort, A. (2007). *Inter-municipal cooperation in Europe*. Dordrecht: Springer.

Kersting, N., & Vetter, A. (2003). *Reforming local government in Europe: Closing the gap between democracy and efficiency*. Opladen: Leske und Budrich.

Kübler, D. & Pagano, M.A. (2012). Urban politics as multilevel analysis. In K. Mossberger, P. John & S.E. Clarke (eds.), *The Oxford handbook of urban politics*. Oxford: Oxford University Press, pp. 114-129.

Ladner, A., Baldersheim, H. & Keuffer, N. (2015). *Self-rule index for local authorities. (Release 1.0)*. Available at:

[http://ec.europa.eu/regional\\_policy/en/information/publications/studies/2015/self-rule-index-for-local-authorities-release-1-0](http://ec.europa.eu/regional_policy/en/information/publications/studies/2015/self-rule-index-for-local-authorities-release-1-0) [Accessed 29 November 2017].

Lefèvre, C. (1998). Metropolitan government and governance in Western countries: a critical review. *International Journal of Urban and Regional Research* 22(1): 9-25.

Lyons, W. E., Lowery, D. & Hoogland DeHoog, R. (1992) *The politics of dissatisfaction: citizens, services and urban institutions*. Armonk: M.E Sharpe.

Mulgan, R. (2014). Accountability Deficits. In M. Bovens, T. Schillemans & R.E. Goodin (eds.), *Oxford Handbook of Public Accountability*. Oxford: Oxford University Press, pp. 545-559.

Neblo, M.A. et al. (2010). Who wants to deliberate - and why? *American Political Science Review*, 104(3): 566-583.

Norris, P. (1999). *Critical citizens. Global support for democratic government*. Oxford: Oxford University Press.

Oliver, J.E. (2012). *Local elections and the politics of small-scale democracy*. Princeton: Princeton University Press.

Papadopoulos, Y., & Warin, P. (2007). Are innovative, participatory and deliberative procedures in policy making democratic and effective? *European Journal of Political Research* 46(4): 445-472.

Perry, S. (2014). Was erwarten Bürgerinnen und Bürger von ihrer Stadt? In J.W. Van Deth (ed.), *Demokratie in der Grossstadt. Ergebnisse des ersten Mannheimer Demokratie Audit*. Wiesbaden: Springer VS, pp. 23-50.

Pierre, J. (2011). Cost-efficiency and public interest in inter-local partnerships. *Urban Research & Practice* 4 (1), 13-22.

Rothstein, B. (2009). Creating political legitimacy. Electoral democracy versus quality of government. *American Behavioral Scientist* 53(3): 311-330.

Scharpf, F.W. (1999). *Governing in Europe: effective and democratic?* Oxford: Oxford University Press.

Schmidt, V.A. (2013). Democracy and legitimacy in the European Union revisited: input, output and 'throughput'. *Political Studies* 61(1): 2-22.

Swianiewicz, P., & Lackowska, M. (2008). Cosmopolitans of small fatherlands. *Miscellanea Geographica* 13: 197-208.

Tyler, T.R. (1994). Governing amid diversity: The effect of fair decisionmaking procedures on the legitimacy of government. *Law and Society Review* 28(4): 809-832.

Tyler, T.R. (2000). Social justice: outcome and procedure. *International Journal of Psychology* 35(2): 117-125.

United Nations (2014). *World urbanization prospects: the 2014 revision*. Available at: <https://esa.un.org/unpd/wup/> [Accessed 29 November 2017].

Warren, M.E. (2014). Accountability and Democracy. In M. Bovens, T. Schillemans & R.E. Goodin (eds.), *Oxford Handbook of Public Accountability*. Oxford: Oxford University Press, pp. 39-54.

Zürn, M., Ecker-Ehrhardt, M. & Binder, M. (2012). International authority and its politicization. *International Theory* 4(1): 69-106.

## TABLES

---

[Introductory Text / Scenario]

Over the next 20 years an increase in traffic is expected for the [X] region. The existing public transport network is reaching its limits.

Imagine that the public authorities in the [X] region have decided to build a new underground line to relieve traffic congestion. To do so, a commission is established to plan this new underground (routing, location of stops) and coordinate construction.

How do you think this commission should be set up? How should it function? We will now present you three times with two possibilities. Please indicate each time, which of the two commissions you would prefer – regardless of whether you support a new underground line in the [X] region or not.

[Attribute]

[Attribute Levels]

The members of the commission...

[Input Legitimacy: Members]

- 1) are directly elected by the citizens of the [X] region [Elected]
- 2) represent the local authorities in the [X] region [Delegates]
- 3) are independent experts [Experts]

Relationship of the commission with the public

[Throughput Legitimacy: Transparency]

- 1) All documents and negotiations are public [Public]
- 2) Only the final decisions will be made public, the negotiations are not open to the public [Not Public]

Cost awareness of the commission

[Output Legitimacy: Budget Over-Run]

- 1) The project budget is exceeded by 0-10% [0-10%]
- 2) The project budget is exceeded by 10-20% [10-20%]
- 3) The project budget is exceeded by 20-30% [20-30%]

The commission makes decisions...

[Formal Authority: Decision Mode]

- 1) by majority vote [Majority]
- 2) unanimously [Unanimous]

The commission's decisions ...

[Formal Authority: Implementation]

- 1) come into force without further agreement of the local authorities in the [X] region [No Approval]
  - 2) only come into force when approved by the local authorities in the [X] region [Approval]
- 

Table 1. Scenario, attributes and attribute levels for conjoint analysis

	Commission A	Commission B
The members of the commission...	Are directly elected by the citizens of the [X] region	Represent the local authorities in the [X] region
The commission makes decisions...	Unanimously	Unanimously
The commission's decisions ...	Only come into force when approved by the local authorities in the [X] region	Come into force without further agreement of the local authorities in the [X] region
Relationship of the commission with the public	All documents and negotiations are public	All documents and negotiations are public
Cost awareness of the commission	The project budget is exceeded by 20-30%	The project budget is exceeded by 0-10%

Which of these two commissions do you prefer? A B

How would you rate **Commission A** on a scale from 1 to 7? 1 indicates that you "don't approve at all" and 7 indicates that you "strongly approve" the commission.

Don't approve at all Strongly approve

1                    2                    3                    4                    5                    6                    7

How would you rate **Commission B** on this scale?

Don't approve at all Strongly approve

1                    2                    3                    4                    5                    6                    7

Table 2. Example of choice task

Country	RAI-Score	LAI-Score	Metro. Area	Capital	Population (2012)	Survey-Respondents		
						Cent.	Surr.	Total
CH	26.5	29.57	Bern	Yes	360,127	193 (35%)	366 (65%)	559 (100%)
			Zurich	No	1,217,751	188 (31%)	419 (69%)	607 (100%)
DE	24	27.24	Berlin	Yes	4,951,687	494 (76%)	158 (24%)	652 (100%)
			Stuttgart	No	2,647,134	153 (25%)	453 (75%)	606 (100%)
FR	10	25.64	Paris	Yes	11,800,687	119 (19%)	522 (81%)	641 (100%)
			Lyon	No	1,934,717	194 (29%)	473 (71%)	667 (100%)
UK	5	17.39	London	Yes	12,208,100	226 (34%)	440 (66%)	666 (100%)
			Birmingham	No	2,873,800	349 (53%)	305 (47%)	654 (100%)
Total						1,916 (38%)	3,136 (62%)	5,052 (100%)

Table 3. Case selection and data

Note: RAI-Score=Regional Authority Index-Score (2010) by Hooghe et al. (2016) (Range: 0-27), LAI-Score=Local Autonomy Index-Score (2010) by Ladner et al. (2015) (Range: 0-37), Cent.=Residents in center city, Surr.=Residents in surrounding area.

# FIGURES

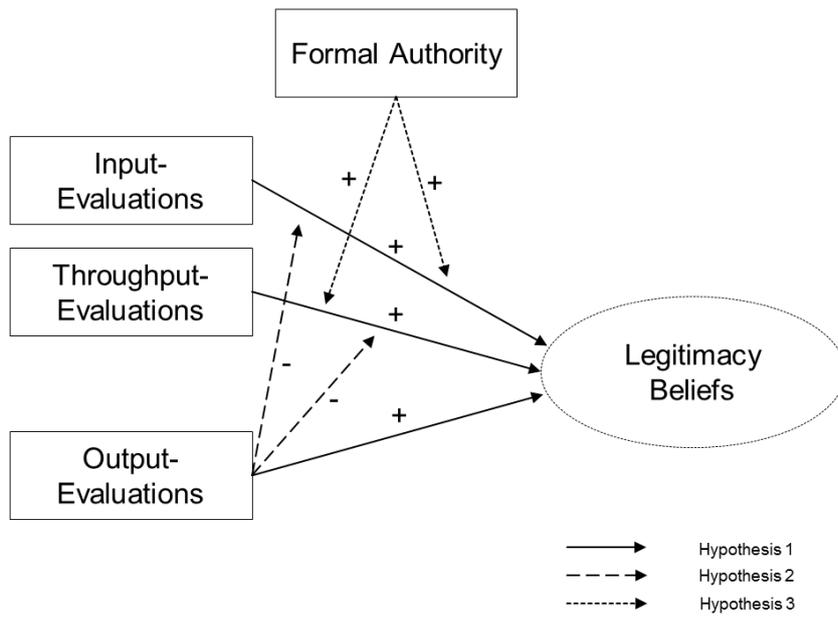


Figure 1. Theoretical expectations

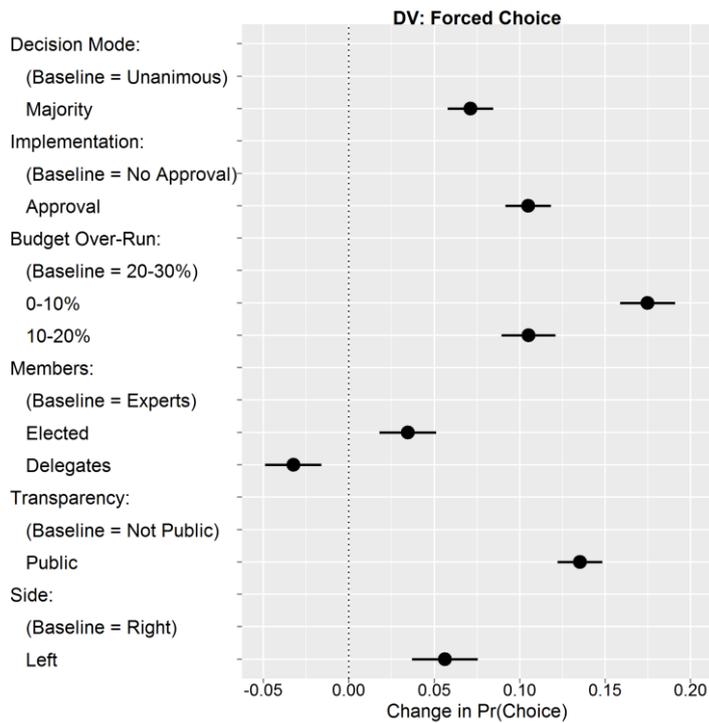


Figure 2. Overall results

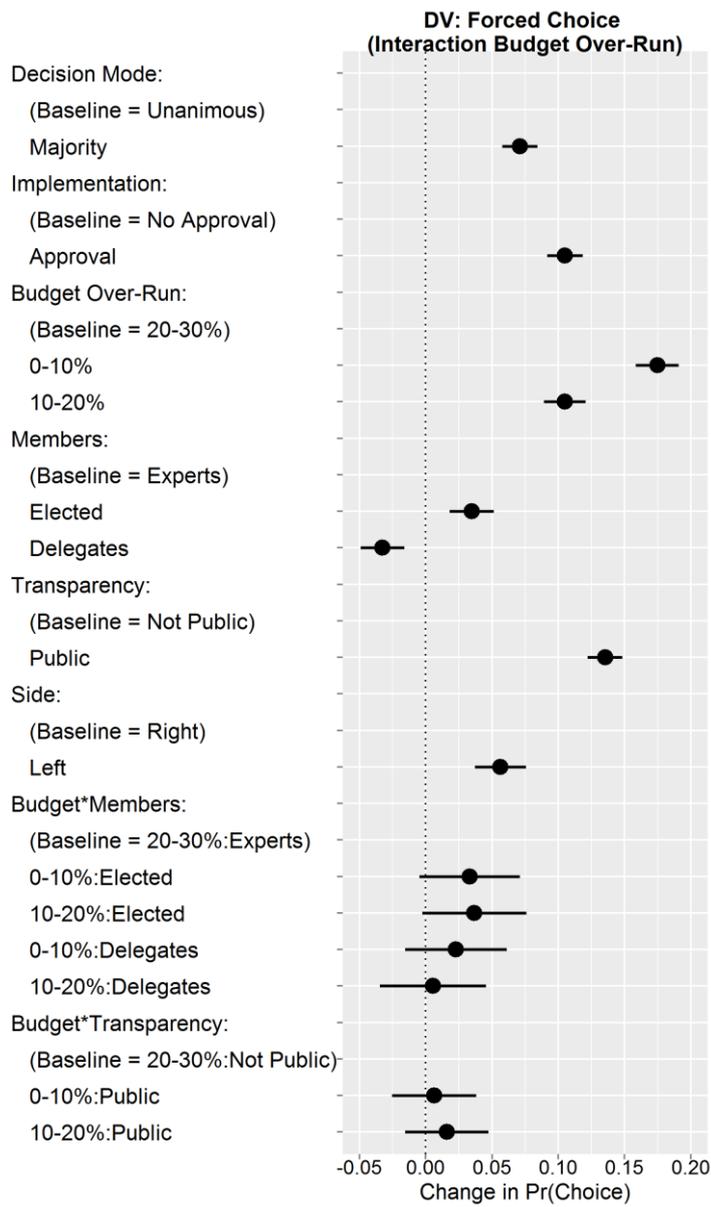


Figure 3. Output interaction: cost efficiency X input/throughput

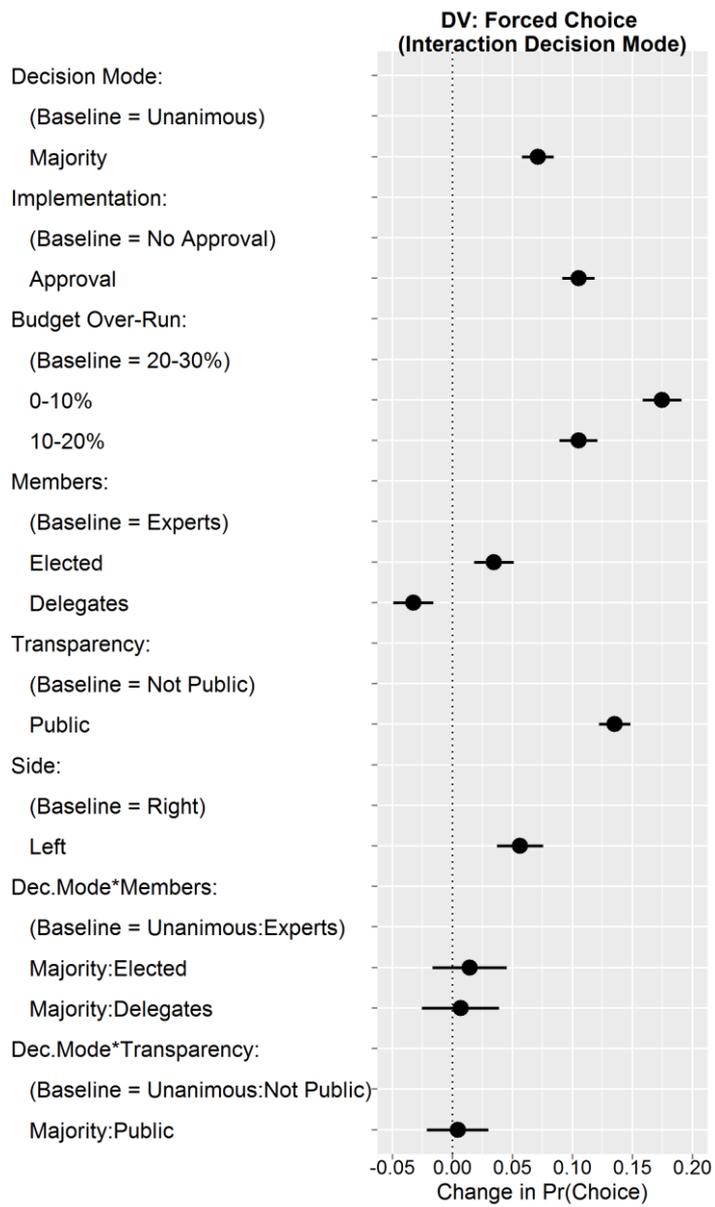


Figure 4. Formal authority interaction: decision mode X input/throughput

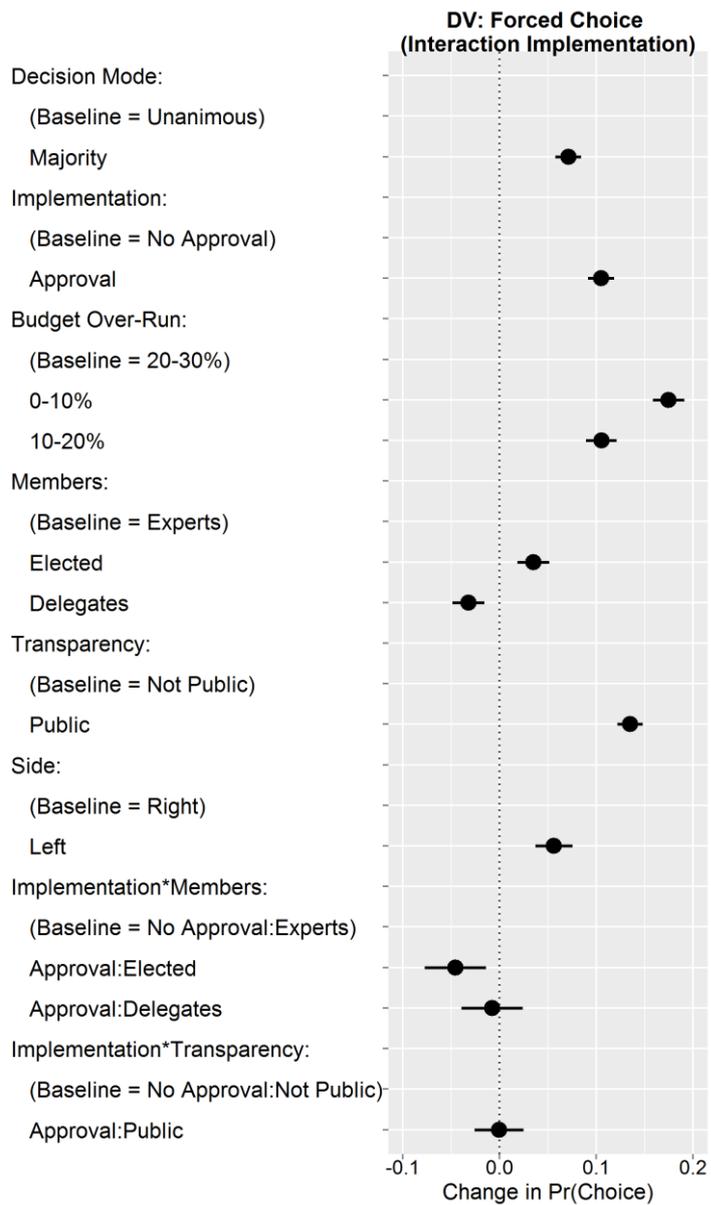


Figure 5. Formal authority interaction: implementation X input/throughput