

THE RESPONSIVENESS OF MPS TO CITIZEN-INITIATED POLICY QUERIES:
PRE-ANALYSIS PLAN¹

Markus Baumann (Goethe University Frankfurt)

Diane Bolet (King's College London)

Rosie Campbell (King's College London)

Tom Louwerse (Leiden University)

Helene Helboe Pedersen (Aarhus University)

Wouter Schakel (Leiden University)

Thomas Zittel (Goethe University Frankfurt)

Abstract

How responsive are legislators to the policy concerns of which kind of constituents? We aim to investigate this question in a comparative correspondence study field experiment (CSFE). In this CSFE, we send unsolicited policy queries via e-mail to all members of the UK *House of Commons*, the Dutch *Tweede Kamer*, the Danish *Folketing*, and the German *Bundestag*, where fictitious senders vary with regard to their ethnicity, gender, class status, and partisanship. With this, we investigate legislators' level of responsiveness to citizen-initiated policy queries and the extent to which this is biased towards the concerns of distinct types of citizens, i.e. partisans and privileged social groups. This paper presents a proposal for a pre-analysis plan for this study. With this study, we contribute to the experimental literature on MPs' individual responsiveness by (a) conducting a comparative CSFEs in four countries, (b) focusing on policy rather than service queries, and (c) including richer measures of the dependent variable (rate, speed and quality of responsiveness).

¹ An earlier version of this pre-analysis plan was presented at the 14th General Conference of the European Consortium for Political Research, 24 – 28 August 2020.

1. Research question and motivation

Responsiveness is a core value in representative democracies. It requires elected officials to take the interests of citizens into account while making policy choices (Pitkin, 1967; Manin, 1997). This involves a number of important empirical implications; the proclivity of individual legislators to personally communicate with individual citizens to learn about their concerns and to provide information and justifications about policy schemes is one of them. This individual responsiveness is found to not only facilitate trust among voters but also interest mobilization as a prerequisite for policy responsiveness and congruence. Consequently, this study asks about the extent to which legislators are responsive to citizens, to which kind of citizens, and why.

Our study advances from previous research that has used correspondence study field experiments (CSFEs) to explore how responsive legislators are to individual constituents and how this is biased by the traits of constituents such as their ethnicity, class status, or gender (Butler and Broockman, 2011; for an overview, see Costa, 2017). This strand of research offers new and innovative ways to address issues of causal inference, e.g. by addressing issues of endogeneity and by exploring distinct forms of behavior in view of counterfactuals. This research also helped to unveil an important empirical detail in the complex process of political responsiveness, i.e. individual level constituency communication. It however remains focused on the American case (for exceptions, see Grohs, Adam and Knill, 2016; Habel and Birch, 2019; Bol *et al.*, 2020), mainly concerned with service responsiveness (for an exception, see Butler, Karpowitz and Pope, 2012), and focused on the incidence of response rather than its content. These are the gaps in this research our study contributes to.

We consider it important to go beyond the service dimension and address the policy dimension of responsiveness. This matters for substantive policies, i.e. what legislators promote in intra- and inter-party debates (Wlezien and Soroka, 2012), and also for the modes of party competition and interest mobilization, i.e. which kind of voters which kind of politicians appeal to, to what extent, and why (Mansbridge, 2003). We also consider it important to expand the geographic scope of this research to better understand how institutional context affects legislators' responsiveness (Heitshusen, Young and Wood, 2005). European democracies offer important variance with regard to electoral and party system contexts that we aim to exploit in this research. Finally, we consider it important to validate the notion of responsiveness in view of e-mail content and thus pay attention to how

legislators respond to individual citizens, whether they indeed provide positions and justifications on the issue that has been raised in the original citizen-initiated policy query.

2. Theoretical frame

When and why are legislators responsive to which kind of citizens? In our study, we advance from an informational model of legislative behavior to provide an analytical lens for comparative research on this issue. This model advances from legislators' electoral motivations. It stresses their efforts to strategically cultivate electorally beneficial contacts in view of distinct cues they take from citizens and contingent upon institutional context bounded by individual-level characteristics.

The need to be strategic about their personal interactions with citizens results from the scarce resources that MPs command while facing a wealth of demands and obligations. We argue that they manage this overload in their representative function by using information heuristics in form of voter cues to selectively respond to citizens (Henderson and Brooks, 2016). In our research we advance from one widely acknowledged approach in this regard that we elaborate below: The partisan constituency approach. Legislators' proclivity to follow this approach is patterned by three key factors: (1) their genuine goals, (2) their institutional context that affects the distinct informational heuristics they adopt, and (3) their own characteristics that may bound their strategic behavior. In the following we briefly elaborate these factors. Again, this model does not aim to provide a comprehensive summary of empirical facts but rather an analytical lens for relevant comparative research.

With regard to legislators' goals, we subscribe to the widely shared assumption, where legislators seek votes, office, and policy, and where these aspirations are hierarchically structured. In this vein, vote-seeking trumps other goals, since it must be seen as a prerequisite to secure office and implement policy (Strøm, 1997). Information heuristics (cues) help MPs to navigate low-information environments and identify those constituents that best facilitate their vote seeking efforts. The established literature highlights one distinct strategy in this regard that we draw from and briefly elaborate in the following: the partisan constituency approach (Ezrow *et al.*, 2010).

The partisan constituency approach assumes that parties align with national coalitions of voters on the basis of common ideological visions. It is said to govern both the behaviors of voters, who provide stable electoral support for those parties they feel closest to, and the behaviors of parties and politicians, who are particularly responsive to partisans (Dalton, 1985; Castles and Wildenmann, 1986; Katz, 1986). The elite-level behavioral effects of the

partisan constituency approach are corroborated by research that has shown that MPs react to partisans and tend to discard the opinions of constituents with whom they politically disagree (Butler and Dynes, 2016). This approach is also corroborated by the high levels of congruence that was found in the policy orientations of partisan voters and their party representatives (Dalton, 1985; Romeijn, 2018).

In our study, we confront the partisan constituency approach with an alternative perspective on how vote-seeking and hard-pressed MPs manage the many demands they face, i.e. what short-cuts they use to decide whom to respond to. In this vein, our informational model assumes, that individual MPs in their personal interactions will react to readily available social references, i.e. pay attention to the visible social characteristics of those individuals that approach them. The social characteristics MPs detect can trigger a variety of mechanisms that translate into responsive/unresponsive behavior. From a strategic perspective, social references help identifying what Boynton, Patterson and Hedlund (1969) portrayed as “attentive constituents” (see also Martin, 2003). Attentive constituents are those who are politically most active and influential and thus most beneficial sources of electoral support. They constitute what V.O. Key (1961, pp. 536–543) called “the thin stratum” lying between mass publics and political elites. In terms of social references, this should put upper-middle class individuals in the spotlight of legislators’ attention, compared to low class voters, since they are commonly found to be more active (Giger, Rosset and Bernauer, 2012; Gilens and Page, 2014; Schakel, Burgoon and Hakhverdian, 2020). Similarly, this also should put ethnic-majority individuals in the spotlight of legislators’ attention, compared to ethnic-minorities, since they have been found to participate less in electoral politics, e.g. to vote and run for public office less.

Figure 1: An Informational model of responsiveness

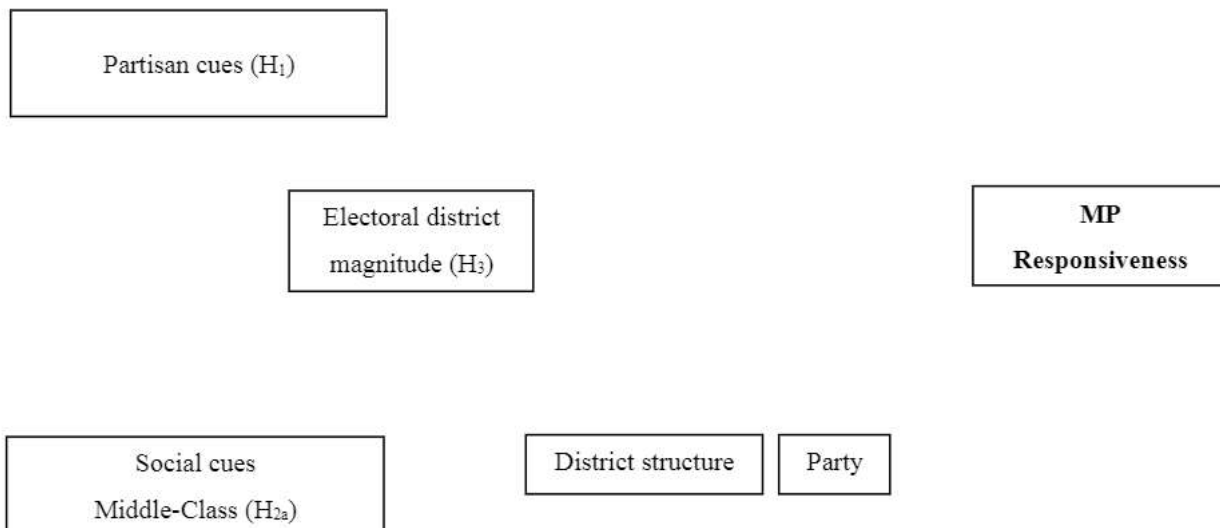


Figure 1 summarizes our overall argument while distinguishing between main effects that we will be able to adequately test in our study and tentative hypotheses that we will follow up in exploratory ways. We elaborate Figure 1 in the following.

With regard to the main effects, we test the two approaches that we outlined above. We however focus on a third main effect since we know from the literature that electoral institutions affect the responsiveness of legislators in direct ways, i.e. their proclivity to engage in personal contacts. This can result from several electoral features. In this study, we highlight the role of district magnitude. Small districts increase the traceability of legislative behavior, facilitates subjective perceptions of accountability to all voters in a geographic district, and also discourages free-rider behavior among MPs representing the same district (Dahl and Tufte, 1973; Mayhew, 1974, p. 87; Cain, Ferejohn and Fiorina, 1987; Cox, 1990).

Beyond our interest in the outlined main effects, we are interested in the relationship between partisan and social cues, and how electoral context, district structure, and also the personal backgrounds of legislators affect their responsiveness, i.e. whether they adopt a partisan-constituency approach or rather use social references to determine whom to respond to. We approach this issue in exploratory ways and refrain from stating testable hypotheses since we cannot be sure that our study provides adequate statistical power for this line of study. We first elaborate our main hypotheses and then get to the exploratory part of our study in the next two sections.

3. Hypotheses

Our theoretical considerations lead us to the following hypotheses that we aim to test in our study.

Advancing from Figure 1, we aim to test the following main hypotheses. First, we advance from the two default models in our argument, i.e. the partisan constituency approach and mobilized voter approach.

- H_1 Legislators are more likely to be responsive to fellow partisans compared to non-partisans

- H_{2a} Legislators are more likely to be responsive to upper middle-class constituents than working class constituents
- H_{2b} Legislators are more likely to be responsive to constituents with majority ethnic backgrounds compared to constituents with a minority ethnic background

Second, we know from the literature that the quantity of legislators' responsiveness depends upon electoral institutions, and particularly upon district size.

- H₃ The smaller the district legislators are elected in, the more likely they are to be responsive.

4. Exploratory analysis

In our study, we are mindful about the contextual circumstances that might affect the quantity and quality of legislators' responsiveness i.e. their proclivity to individually respond and whether they follow a partisan-constituency or an attentive-voter approach. We particularly focus on the following five issues that we depict in Figure 1.

First, party might matter as a result of different mechanisms. For example, traditional cleavage theories stress stable alignments between distinct parties and distinct social groups. In this vein, members of left parties should be more inclined to represent constituents with low class and minority backgrounds compared to MPs representing right parties, independent of partisan cues. Also, populist parties claim to promote responsiveness and to criticize established parties for being too distant from the average voter. As a consequence, we would expect MPs representing populist parties to be more responsive than the rest of their colleagues, independent of social cues and partisan cues. Lastly, party size might matter. Small parties might be less able to provide practical help and support to their MPs. As a consequence, the members of small parties might be less responsive, independent of whom they approached by.

Second, district structure. MPs that face significant segments of minority-ethnic voters or low-class voters should be more inclined to represent constituents with low class and minority backgrounds, independent of partisan cues. In such districts, vote-seeking MPs should wish to make sure to not alienate major segments in their constituency and prevent disappointed individuals who did not receive a response from spreading the word. Also, with regard to the personal perceptions of MPs, taste-based discrimination of lower class or ethnic-

minority constituents should be less likely in districts that involve large segments of these groups.

Third, district magnitude. MPs elected in small districts should be more inclined to respond to a broader segment of their constituency, i.e. adopt an attentive-voter approach, since they face a smaller number of constituents, enjoy a more direct relationship with them, and also are to a lesser extent subject to free rider concerns (Norris, 2004; Carey, 2007; Miler, 2010). MPs in large or even national constituencies need to construct their electoral coalition in view of a complex electoral context. This should render them most likely to resort to partisanship as a most efficient way to manage information overload.

Fourth, the personal background of MPs. MPs should be more inclined to represent constituents who share a similar background, independent of electoral considerations, i.e. partisan cues or electorally viable social references. That is to say, majority-ethnic MPs are expected to be more responsive to majority-ethnic citizens, and minority-ethnic MPs are more responsive to minority-ethnic citizens. Similarly, we expect that MPs with academic educational backgrounds are more responsive to upper middle-class citizens, while MPs with non-academic backgrounds should be more responsive to working-class citizens. This results from what is referred to in the sociological literature as “homophily”, where individuals tend to be drawn to those that mirror their traits and characteristics (Schneider *et al.*, 1998; McPherson, Smith-Lovin and Cook, 2001). But this also meets the expectations of prescriptive analyses on the issue of descriptive representation (e.g. Mansbridge, 1999). It furthermore has been corroborated in previous experimental research on the issue (Broockman, 2013; Butler, 2014) that however has largely been confined to the American case.

Fifth, we are interested in the interaction between partisan and social cues. Our factorial design allows to explore whether the absence of partisan cues exacerbates the effects of social cues, i.e. whether legislators are particularly inclined to adopt an attentive voter approach if they face non-partisan constituents. In view of weakening partisanship in Western democracies, this question is of key concern.

We do not develop hypotheses on related effects that we aim to test. We rather will use them retrospectively to explore our experimental data and to conduct follow up analyses employing additional methods such as quantitative text analysis with observational data. This results from the fact that we cannot be sure that our study provides adequate statistical power for this line of study. This is not special to our study but rather a common issue with interaction effects. For example, Gelman (2018) has shown that one would need 16 times the

sample size to detect an interaction than to detect a main effect when the interactions are half the size of the main effects.

5. General research design

To test the above-mentioned main hypotheses, we conduct a correspondence study field experiment (CSFE) in Denmark, Germany, the Netherlands and the United Kingdom, scheduled for November and December 2020.

In our CSFE, emails written by fictitious constituents are sent to inquire about the positions of MPs on distinct policies. We vary sender characteristics, i.e. name unveiling gender and ethnicity, stated occupation, and information about party identification. With this, we test the effect of informational heuristics on the quantity and quality of responses among MPs, i.e. the effect of partisan and social cues.

Our comparative design involves four European parliamentary democracies that all fall into the category of strong party government. Though there are various differences in institutional settings, one major difference is found in their electoral systems. The United Kingdom has a plurality system that establishes a direct accountability link between individual legislators and (geographic) constituents in single member districts, where district size equals “1”. On the other side of the spectrum, elections in the Netherlands are governed on the basis of a proportional formula in a de-facto national multi-member district, where district size equals 150. The Danish and German cases provide for intra-case difference in district size, even though to different degrees. Danish elections are contested on the basis of a proportional formula in 10 multi-member districts that range between 2 and 20 seats. The German mixed-member proportional system deliberately aims to combine a partisan model of representation with some form of personalized accountability. It results in a set of MPs that is elected in a plurality vote in single member districts (N=299) and a set of MPs elected in 16 multi-member districts on the basis of a proportional formula (currently N=408). The multi-member districts vary in size between 4 and 134.

In short, our comparative design allows us to explore the effects of informational cues on legislators’ responsiveness across different electoral rules, i.e. across differences in district size. Our Small-N design involves potential third variable biases at the country level that we cannot entirely control for. But it allows to observe patterns that can be matched with other quantitative and qualitative data and that allow to dig deeper in further research. It

furthermore involves a fair share of variance in the Danish and German cases that allow to test the effects of district size independent of country level variables.

In terms of its implementation, our field experiment differs from most other correspondence studies in several ways. One such difference is that we use a factorial design. That is to say, while most previous studies manipulate one variable at a time, such as the sender's ethnic background, we manipulate several variables in the same emails. There are several reasons for this. First, a factorial design allows us to compare the effects of multiple variables without sending out many waves of emails to MPs. Second, there is no inherent trade-off between the number of factors (variables) in the treatment and statistical power, as the effect of one factor can be interpreted as the average marginal effect over all other factors (Gerber and Green, 2012, pp. 395–400).² If anything, this increases the generalizability of our findings.³ An important caveat here is that we are very limited in our ability to detect interactions between factors, particularly higher order interactions, but this is not necessary to test our main hypotheses. Third, there is some information that is conveyed in the email by default, particularly the gender and ethnic background of the sender, which is signaled by their name. Hence, we can vary these characteristics without negatively impacting the length, complexity or realism of the email.

The factors we vary in the treatment emails are partisanship, social class, ethnic background, and gender. As we explained above, three of these four cues correspond to distinct informational approaches in legislators' constituency communication. We do not have any expectations about the effect of citizens' gender, but we vary this because it is signaled either way by the sender's name, as we mentioned above. By sending half of all emails from women and half from men, the effects of the other independent variables represent the average marginal effect over a realistic gender distribution (Muralidharan, Romero and Wüthrich, 2020, p. 31). We discuss the operationalization of our independent and dependent variables in section 6 below.

Beyond its factorial design, our field experiment differs from previous correspondence studies in the sense that we intend to send out two waves of emails to the same MPs. In other

² While there is no inherent trade-off, there are some conditions under which the inclusion of one factor may introduce attenuation bias for another factor. For one thing, this may occur when the treatment becomes overly long and complex, and any individual factor no longer stands out (Bansak *et al.*, 2019). Furthermore, if some factors are sufficient conditions for the outcome variables, this makes it impossible for any other factors to have much of an effect. However, neither of these conditions are likely to apply in our case.

³ For example, in a 2x2 design, where the ethnicity and occupation of the sender are varied (Habel and Birch, 2019), the marginal effect of occupation is really the average of the marginal effect of occupation for ethnic minority and ethnic majority citizens (Muralidharan, Romero and Wüthrich, 2020). If only occupation is varied and ethnicity is kept constant – with all emails coming from purported majority citizens – the marginal effect of occupation only reflects its effect for ethnic majority citizens.

words, we use a mixed design that has both a between-subject component and a within-subject component, in a similar way to conjoint experiments (Hainmueller, Hopkins and Yamamoto, 2014). The reason to field multiple waves of emails is to increase the statistical power of the experiment. While we collect data for the entire population of MPs in Denmark, Germany, the Netherlands and the United Kingdom, our data can still be regarded as one realization of many measurements from a hypothetical superpopulation. Moreover, we aim to generalize beyond the four countries to West European societies with similar electoral systems. For these reasons, the threshold of statistical significance is still relevant for our study. We elaborate on the analysis in sections 8 and 9.

6. Operationalization of variables

As alluded to above, our main independent variables include social class, ethnic background and partisanship, which are all manipulated in the treatment emails (see section 8), in addition to gender. We will discuss the operationalization of each of these variables in turn.

First, social class is a concept whose meaning and measurement has long divided scholars. In line with conventional approaches, however, we view occupation as the best single indicator of social class (Erikson and Goldthorpe, 1992; Oesch, 2006; Carnes, 2013). A person's occupation tells us something about their income, wealth, educational background, job prestige and social network. Occupation also has a practical advantage over indicators like income or wealth due to the fact that it is arguably more artificial to signal the latter in an email to an MP. This presumably explains why previous correspondence studies have also relied on occupation (Butler, 2014; Carnes and Holbein, 2019; Habel and Birch, 2019).

For our treatments, we select two occupations that are commonly regarded as belonging to the lower (or 'working') class and upper middle class, respectively. We base this on Harry Ganzeboom's index of socioeconomic status, which classifies all occupations in the International Standard Classification of Occupations (ISCO) on the basis of their income, education and job prestige (Ganzeboom, De Graaf and Treiman, 1992). We choose two occupations that are close to either end of the status scale. Moreover, we select jobs that are relatively common in all three of our countries, to ensure that MPs know which social class

the job belongs to and also to reduce the risk of exposure.⁴ Based on these criteria, we use a *cleaner* as a working class occupation and a *lawyer* as an upper middle class occupation.

Next, we vary the ethnic background and gender, which are both signaled by the name of the sender and will therefore be discussed together. Like social class, the concepts of ethnicity and gender are subject to heated debate and evolving definitions. For our purposes, these debates are of substantive interest, as we choose names that are commonly regarded as being typically male or female, or in the case of ethnicity, belonging to native-born citizens or people with a ‘non-Western’ background. Hence, we select first names that are among the most common names given to men and women born between 1950 and 1990.⁵ We match these to the most common surnames in the same time period. For ethnic minorities, we choose first and last names that are among the most common among citizens with a ‘non-Western’ migration background.⁶

As the above suggests, we use several names within each category (majority male, majority female, minority male and minority female). This is done to “[make] results less dependent on idiosyncratic design decisions” (Gerber and Green, 2012, p. 398), such as the possibility that a specific name ends up in the news when our experiment is in the field. The specific names selected for the field experiment are listed in Table 1.

We have registered email accounts under these names with Gmail in October of 2020. For the protocol of registering said accounts, see appendix A.

Lastly, partisanship is manipulated in a straightforward way by either stating that the sender generally supports the MP’s party or not including any indication of his or her partisanship. This sends a cue to the MP about their likely electoral value on the basis of party-based ties.

In addition to these above-mentioned independent variables, we are interested in the role of district size, district structure, party, and legislators’ personal characteristics, i.e. their ethnic and class backgrounds. These are factors that we do not manipulate but only measure.

⁴ As mentioned above, multiple emails will be sent to each MP. If an MP receives several emails from people with an unusual occupation, this may arouse suspicion and alert them to the nature of the field experiment.

⁵ This window is chosen to exclude names that are mostly found among older or younger people, as we want to avoid associating a name with a certain age group. In reality, however, the relative popularity of names is quite stable over time.

⁶ For the Netherlands, these are inhabitants with a Turkish background, who numbered 418,869 people, respectively, as of July 2020. For Germany, these are also inhabitants with Turkish backgrounds. This group is with around 2.82 Mio. (2019) by far the largest ethnic group in Germany, ahead of inhabitants with Polish backgrounds (2.23 Mio.) and with Russian backgrounds (1.39 Mio.). The largest ethnic groups in the UK are Indian (2.5 per cent) and Pakistani (2 per cent). However, the most common male name is of Arabic descent, Mohammad, hence the choice of Arabic names and surnames for both genders to make the treatments comparable. In Denmark the largest minority group among inhabitants in the fall 2020 is also Turkish (1.1 percent, 64,341) followed by Poles (0.8 percent, 48,213) and Syrians (0.8 percent 43,443).

Given the small number of countries, district size is a factor that is potentially confounded by other possible country differences, i.e. any interpretation of differences will reflect the fact that the risk of omitted variable bias is as large as it is in observational studies. We nevertheless operationalize this electoral factor at the individual level and measure the logged size of the district in which a legislator has been elected.

Table 1. Names of fictitious constituents in the field experiment

Gender	Ethnicity	Names			
		Denmark	Germany	Netherlands	United Kingdom
Female	Majority	Anne Nielsen	Melanie Wagner	Ingrid Bakker	Jessica Smith
Female	Majority	Anne Hansen	Melanie Becker	Ingrid van Dijk	Jessica Jones
Female	Majority	Kirsten Nielsen	Julia Wagner	Linda Bakker	Emma Smith
Female	Majority	Kirsten Hansen	Julia Becker	Linda van Dijk	Emma Jones
Male	Majority	Peter Nielsen	Christian Wagner	Jeroen Bakker	Thomas Smith
Male	Majority	Peter Hansen	Christian Becker	Jeroen van Dijk	Thomas Jones
Male	Majority	Jens Nielsen	Stefan Wagner	Willem Bakker	Paul Smith
Male	Majority	Jens Hansen	Stefan Becker	Willem van Dijk	Paul Jones
Female	Minority	Fatma Yilmaz	Aylin Yilmaz	Fatma Yilmaz	Yasmin Khan
Female	Minority	Fatma Celik	Aylin Kaya	Fatma Demir	Yasmin Hassan
Female	Minority	Ayse Yilmaz	Elif Yilmaz	Selma Yilmaz	Maryam Khan
Female	Minority	Ayse Celik	Elif Kaya	Selma Demir	Maryam Hassan
Male	Minority	Mehmet Yilmaz	Malik Yilmaz	Muhammed Yilmaz	Mohammad Khan
Male	Minority	Mehmet Celik	Malik Kaya	Muhammed Demir	Mohammad Hassan
Male	Minority	Mustafa Yilmaz	Amir Yilmaz	Ibrahim Yilmaz	Ali Khan
Male	Minority	Mustafa Celik	Amir Kaya	Ibrahim Demir	Ali Hassan

As we outlined in the above, we are furthermore interested in the role of district structure and party. Again, these are factors that we do not manipulate but measure. With regard to district structure we use available statistical information to particularly tap into the economic situation of district inhabitants, i.e. the share of lower-class constituents, and their ethnic composition, i.e. the share of constituents with ethnic-minority background. With regard to parties, we distinguish between left and right parties on the basis of the left-right variable (“Irgen”) from the 2019 Chapel Hill Expert Survey.

We are furthermore interested in legislators’ personal backgrounds and distinguish between legislators with upper-middle class and lower-class backgrounds and between legislators with minority and majority ethnic background. For this purpose, we collect data on MPs ethnic background (measured on the basis of self-reported and/or third-party information) and education (measured as a dichotomous variable that distinguishes MPs with and without academic education, i.e. a bachelor degree or higher).

Regarding our *dependent variables*, we measure several outcomes. The first and most obvious is whether the MP sent a reply to our fictitious constituent. This excludes automatic replies. In the unlikely event that an MP cannot be reached – that is, if the email is undeliverable – we will resend the email the next day and, if necessary, try to find an alternative email address. If the MP cannot be reached on any email address, we consider them non-responsive. As a robustness check, we delete them from our analysis.

Since the presence or lack of a reply is only a partial proxy of personal responsiveness, we also measure other features of a potential response. Our second outcome variable is the time it takes for the MP to reply, measured in 24-hour periods after sending the email (excluding weekends). We collect responses for four weeks after the moment the email is sent. If an MP has not responded by then, we code this as a lack of a response. A third outcome is the length of the reply in words, using a logged word count. Lastly, we manually code whether each reply is responsive in the sense that it answers the substantive question which was posed by the sender with references to relevant positions, activities, and information (excluding attached or linked party manifestos and materials). The coding instructions for this last dependent variable are provided in appendix B below.

To prevent extremely skewed distributions that mainly reflect whether a response was sent at all, the second, third and fourth outcomes are dichotomized. For the speed of the reply, all MPs who respond faster than the overall median response time (including non-responses) are given a 1, while the others get a 0. Similarly, emails with more words than the median word count get a 1; shorter emails are scored as 0.⁷ And the ‘informativeness’ of the response is also coded as a dichotomous measure. To avoid post-treatment bias, non-responses are given a 0 on all outcome variables (Coppock, 2019). Our main dependent variable combines all four outcomes into an additive index, where all indicators are equally weighted.⁸

It is important to add that we will use item response theory scaling methods to test whether our four dependent variables are actually part of a single underlying dimension. If this is not the case, we will adjust accordingly. For example, in the plausible case that the speed of a response does not strongly correlate with the quality or the length of the response, we will remove this from the index. If the scaling method does reveal the presence of a single dimension of responsiveness, however, we will use our additive index as the main dependent

⁷ To be clear, this is the median word count among all replies, so excluding non-responses. To account for the possibility that some languages are inherently briefer than others, we use data on translated speeches in the European Parliament between 1996 and 2011 (Koehn, 2005). This corpus shows that Danish and German are slightly briefer than English while Dutch is less brief, at least in terms of political speech. Based on this data, we will multiply the overall median word count by 0.919 in Denmark, by 0.932 in Germany and by 1.023 in the Netherlands. This gives us adjusted cutoff points for short and long replies in all countries, while still allowing for between-country differences in responsiveness.

⁸ In other words, MPs can score a maximum of four points, which they can get by sending a fast, elaborate and informative response. We expect the four indicators to be strongly positively correlated, but even if they are not, this does not necessarily undermine the validity of our index. For instance, in interviews with former MPs, some noted a trade-off where an MP can either send a brief but prompt reply or wait longer and send a more elaborate reply. If this applies, the second and third outcomes will be negatively or weakly correlated. However, we do not want to judge either of these response types as being better than the other and hence we weight all indicators equally.

variables. As a robustness check, we will also conduct analyses using the separate indicators as dependent variables (see section 9 below).

7. Sampling and assignment strategy

Our sampling strategy is straightforward, given that the experiment includes (almost) all current MPs in Denmark, Germany, the Netherlands and the United Kingdom (as of November 2020). The sample therefore includes 179 MPs in Denmark, 706 MPs in Germany, 150 MPs in the Netherlands and 648 MPs in the United Kingdom.⁹

The assignment of MPs to treatment conditions is a bit more complicated, given the factorial and repeated nature of our design. First, each of our factors has two levels, which means we have a total of $2^4 = 16$ conditions. These 16 conditions are displayed in Table 2 below.

Table 2. Treatment conditions in a 2×4 factorial experiment

		Ethnic majority		Ethnic minority	
		Working class	Upper middle class	Working class	Upper middle class
Female	Partisan	1	2	3	4
	Independent	5	6	7	8
Male	Partisan	9	10	11	12
	Independent	13	14	15	16

The 16 conditions are assigned randomly to MPs within each wave. The major exception to this is that an MP cannot be assigned to the same condition more than once, nor can they receive two emails from the same citizen. Furthermore, there are a small number of MPs (14 in all four countries combined) who have split off from their party in the current legislative term and are now independent representatives. We exclude these from the ‘partisan’ treatments, since a citizen cannot tell them they support the MP’s party.

To ensure balanced treatment conditions and improve statistical precision with our limited number of participants, we use block random assignment based on the seat share and government-opposition status of political parties, as well as logged district size (Gerber and Green, 2012, pp. 71–80). Simply put, MPs are put into subgroups based on these characteristics, and random assignment takes places within each of these subgroups. Since

⁹ Three German MPs are excluded because they left parliament shortly before the start of fieldwork. Two British MPs are excluded because they do not wish to be contacted via email.

this is done separately for each country, district size is only relevant in those countries where it varies among MPs (Denmark and Germany). The blocking procedure is carried out using the `blockTools` package in R.

One further nuance is added to this: we use multiple versions of the treatment email, where the same topic is addressed but using different wordings each time. In total, we have two different email versions, corresponding to the number of waves. Each MP will receive all two versions in a random order, meaning they never receive the same wording twice. By design, this assignment is orthogonal to the factors in Table 2. We believe this is necessary to reduce the risk of exposure, which poses a real danger to the field experiment as a whole.¹⁰ Put differently, this is a likely source of imprecision which is necessary to avoid a much bigger source of bias (Butler and Crabtree, 2017). We do not assume that the different versions of the email are equivalent and will control for this variable as a robustness test (see section 9). Essentially, this can be regarded as another factor with three levels, making for a total of $2^4 * 2^1 = 32$ conditions.¹¹ The only difference with the other factors is that we are not interested in the substantive difference between its levels, nor do we have any priors about this.

Emails are scheduled to be sent out in a period of four weeks, starting in November of 2020. In Denmark, fieldwork is delayed until December and January of 2020 due to practical limitations, but the procedure is otherwise the same. MPs are randomly assigned to receive the first email in either the first or the second week of our fieldwork period. This assignment is done within parties, such that MPs of the same party are spread out over both weeks. The former group will receive the second wave of emails in the third week, while the latter group will receive their second email in the fourth and final week of fieldwork. This ensures that there is at least one full week (and two weekends) between the two waves for all MPs. Within each week, MP are assigned randomly to a week day. Emails are sent out between 8:00 and 20:00 on each working day (local time).

The full code that was used to assign MPs to treatment conditions and waves is provided with this preregistration plan.

¹⁰ If MPs express a belief that the email was not sent by a real person, their response data is unusable. Moreover, other MPs could well believe the same thing and, even though they may not tell us as much, this would still influence their response. In the worst case, news of the experiment spreads among MPs, in which case the response of one MP will be influenced by the treatment received by another MP. In this way, exposure threatens both of the core assumptions behind experimental research, namely that of excludability and non-interference (Gerber and Green, 2012).

¹¹ We have not displayed all 32 of these conditions in Table 2 to keep this table legible.

8. Text of treatment emails

In our treatment emails, we ask MPs a policy-related question. This differs from most other correspondence studies, which usually ask service-related questions, such as how to register to vote (Butler and Broockman, 2011). Our choice is based on the fact that inquiring about an MP's policy stance – and MPs explaining their stance to voters – is an important dimension of responsiveness, as we explained above.

Despite its importance, the choice of a policy-related question raises challenges that a service-related question does not, which may explain why the latter has been more common. First, our emails should address a policy issue that is equally relevant to all political parties, so as to avoid biasing our results against specific party types or ideological leanings. Second, the policy issue should be salient and thus of equal importance to all MPs, mainly to prevent MP's from all too easily forwarding the email to another member of their party who is the policy specialist in the relevant area. This is particularly likely to happen in parliaments where parties are very much based on policy specialization. While this is the case in all modern parliaments, this might be even more so in smaller parliaments and/or in parliaments with many parties, i.e. in the Dutch and the Danish case. Given that this would violate the non-interference assumption and greatly increase the risk of exposure, we want to avoid this. For that reason, we select an issue that is salient enough to touch on all major policy areas.

A third and related consideration is that we should not state or signal any policy views in the email, but should instead ask an open-ended question. The reason for this is that we clearly want to avoid creating the illusion that there is support for a specific view among citizens when there may not be, as this would be highly unethical. However, even expressing concern about an issue in the lead-up to a question sends a signal to MPs about the public agenda. To deal with this, we come back to the first and second factors: by raising an issue that is relevant to all parties – which they are free to address from their specific angle – and raising an issue that is known to be salient among the general public, we avoid distorting the link between citizens and elites.

We believe that, given the current political and public agenda (as of October 2020), the issue that best fits these criteria is COVID-19, and particularly its effects on economic and social life. This is clearly among the most, if not the most, salient issues in all four countries. It also allows us to ask credible questions, as there is much uncertainty among citizens about the topic. It touches on almost all political portfolios, and all parties across the ideological spectrum have taken positions on the topic. For instance, some have argued that the corona

crisis has reinforced the importance of generous welfare state measures, sustainable growth and social cohesion, while others stress the need for supply-side stimulus, budget-neutral fiscal policy and personal responsibility. Our treatment does not preclude any of these responses.

Below, we present the email templates that will be sent to MPs, where treatment levels are written in brackets. Danish, German and Dutch translations of these emails are presented in appendix C. It should be pointed out that the templates are the same in all countries, with one exception: the British and German versions version include a sentence, or part of a sentence, where the citizen states that (s)he has recently moved to the MP's constituency. This is necessary since it diminishes exposure risk, especially in Germany where MPs might forward out-of-district mails to policy experts in their party group, and since it also secures the credibility of the mails in setting that involve single member districts.

Version 1

Email subject: Question about Covid-19

Dear [Title] [MP's surname],

My name is [first name] [surname] {and I've recently moved to the constituency you represent.} I'm working as a [cleaner / lawyer] and I'm worried about the consequences of the Covid-19 crisis. I work for a large company and I personally feel safe for now. But I'm worried about the longer term. I see the crisis affects people all around me who are losing jobs or experiencing pay cuts. And many other problems are being neglected because everything is about corona now.

[As a [party name] supporter / statement left out for control group,] I'd like to know what are you and [party name] are going to do to get us through this crisis in the best possible way.

I am looking forward to your response.

Best wishes,

[Constituent's name]

Version 2

Email subject: Query about coronavirus

Dear [Title] [MP's surname],

I am [first name] [surname]. {Recently, I've moved to your constituency.} I'm emailing you because I am concerned about the impact of the corona crisis. I am working as a [cleaner / lawyer] and I see a lot of people around me who are suffering as a result of the crisis, losing their jobs or facing pay cuts. As I work for a big firm, I am safe for now. But I am worried about the future. I feel anxious not only because of corona specifically, but also because all the other problems don't get much attention because of corona.

I would like to ask you [,as a [party name] supporter / statement left out for control group,] how you and [party name] are planning to guide us through these difficult times.

I am looking forward to your answer.

Kind regards,

[Constituent's name]

9. Analysis plan

To test our hypotheses, we work with a pooled data set. We analyze these data in view of several models. Our main model tests for differences-in-means using an ordinary least squares estimator with heteroskedasticity-robust standard errors, and takes the following form, using partisanship as an example:

$$Y_i = \beta_0 + \beta_1 \text{Partisanship} + \epsilon_i$$

where Y_i represents the outcome Y for observation i – more specifically, the response index described in section 6 (conditional on its validation) – and β_0 represents the intercept. The coefficient β_1 represents the key causal parameter, tested using a dichotomous variable

for the treatment factor we manipulate in the experiment. Our hypothesis is that $\beta_1 > 0$, which we test in one-sided tests (given the clear direction in the expected effects) with $\alpha = 0.05$. Other treatment factors (ethnicity and class) are tested in the same way but do not have to be included in the same model since they are assigned independently of each other.

Observations are email waves nested in MPs. Different observations within each MP cannot be regarded as independent of each other (see section 10). For this reason, we cluster standard errors by MP in all analyses, unless otherwise specified. In all models, we apply country-level weights so that all four countries are equally influential and the results are not only driven by Germany and the United Kingdom, which involve by far the greatest numbers of MPs.

Our baseline analysis excludes MPs who respond to the treatment email by only asking for a postal address or follow-up conversation (e.g. a phone call instead of an email). This is particularly likely to happen in Great Britain. We remove these MPs as we do not send follow-up emails to MPs (see below), and in the absence of this, there is too little information to label them as either responsive or non-responsive. Since such requests for an address or follow-up conversation are unlikely to be correlated with the treatment conditions, we assume this will not cause bias and will only result in a loss of precision.

In addition to this main model, we will conduct exploratory analyses. We remain cautious about this for reasons explained above. In these exploratory analyses, we add variables that tap into party ideology, district structure, electoral context, and legislators' characteristics and interact them with β_1 , β_2 and β_3 . For the sake of clarity, we will estimate separate models for each of the three treatment factors and for each of the four sets of variables. To provide an example, the model for our partisanship treatment that we interact with electoral context takes the following form. Again, we use one-sided tests with $\alpha = 0.05$.

$$Y_i = \beta_0 + \beta_1 \text{Partisanship} + \beta_2 \text{DistrictSize} + \beta_3 \text{PartisanshipDistrictSize} + \epsilon_i$$

In addition to this, we will conduct a number of analyses to test the robustness of our findings for our main models. First, as mentioned above, we will use alternative dependent variables which measure (a) whether the MP responded to the treatment email at all, (b) how quickly they responded, (c) how long their response was and (d) whether this response was informative. Second, we will use fixed effects for MPs instead of clustered standard errors.

Third, we will use ordinal logistic regression to analyze the responsiveness index instead of OLS. Fourth, and although this is unlikely to matter, we will add a number of control variables to check for covariate balance and possibly improve the precision of the causal estimates. These variables include dummies for the version of the treatment email (see section 8 above), country dummies, and MPs' party. Fifth, we will control for order and period effects by including dummies for the order in which MPs received the emails and dummies for each email wave. Sixth, we will add a dummy variable for the 14 independent MPs (see section 7 above) to ensure they do not affect covariate balance. Seventh, we will check that responsiveness does not differ between the different constituent names (see section 6 above) by testing the explained variance of a model with dummies for majority female, majority male, minority female and minority male citizens versus the explained variance of a model with dummies for all different names. Eighth and finally, we exclude MPs who could not be reached on any email address (see section 6 above).

10. Power analysis

To calculate the statistical power of our envisioned analyses, we rely on data simulations, which are more flexible than standard power calculators given the repeated nature of our treatment (full replication code is provided in the attached files). As with any power analysis, we need to have estimates of some parameters, which in this case include the effect sizes, the average response rate of MPs in each of the four countries and the share of the variation in responsiveness that is within and between MPs.

To simplify the analysis, we use a dichotomous measure of responsiveness; that is, whether the MP responded to the treatment email or not. As mentioned above, this is not our only dependent variable, but we use it here because it is common in previous CSFEs and hence allows us to make informed estimates of all main parameters. A further consideration is that any effect of our treatment factors is likely to be larger for the full response index than for the dichotomous responsiveness measure. After all, the latter is more refined than the former, and it is likely that any bias in the length, timeliness and informativeness of a response works in the same direction as a bias in whether a response is sent at all. Hence, this choice will make for conservative estimates of statistical power.

The effect size of the treatment factors is hard to estimate, in part because most previous CSFEs are limited to the United States. With this geographic bias in mind, Costa (2017) reports an average effect of almost ten percentage points for ethnicity in her meta-analysis.

Social class has been explored less often than ethnicity, but some previous studies find quite small effects here, of two percentage points or less (Carnes and Holbein, 2019; Habel and Birch, 2019; cf. Butler, 2014). Regarding partisanship, Butler and Broockman (2011) report an effect of 4.5 percentage points. Based on these very partial estimates, we consider an expected effect of five percentage points to be a reasonable estimate for our treatment factors.

To estimate average response rates in all four countries, we can also partially rely on previous CSFEs. The study of Bol et al. (2020) reports a response rate of 63% for Germany, while Magni and Ponce de Leon (2020) find a response rate of 53%, and hence we set the expected response rate for the German case to 58%. Habel and Birch (2019) find a remarkably high response rate in the United Kingdom of around 90% that they explain with parliamentary rules that require MPs to respond to constituents. However, a website that encourages citizens to write emails to their MPs and collects statistics regarding their responses shows a much lower response rate of roughly 50%.¹² Based on this, we assume an expected response rate of 80% in the United Kingdom, where the previous CSFEs weighs more heavily in determining our priors as it is similar to our planned experiment. For the Netherlands, we rely on the only CSFE that has been conducted here (to our knowledge), which found a response rate of 42% (Magni and Ponce de Leon, 2020). To our knowledge, there have been no previous CSFEs in Denmark, making the response rate hard to predict. For this reason, we simply assume a response rate of 50% here.

The most difficult parameter to estimate is the share of the variation in responsiveness that is ‘within’ MPs rather than between them. This affects the extent to which sending out multiple waves of emails actually increases statistical power. We do not know of any previous CSFEs that have used a within-subject design and hence we turn to other sources for this. First, the aforementioned website that encourages British citizens to write emails to their MPs reports the response rate of each MP, and using this information produces an intraclass correlation of 0.15 (with email waves as the level 1 variable and MPs as the level 2 variable). Second, self-reported response rates from a 2019 survey of Dutch MPs lead to an intraclass correlation of 0.27.¹³ Third, we can use information from a German website that allows citizens to publicly ask questions and potentially receive answers from their MPs.¹⁴ The

¹² <https://www.writetothem.com/stats/2015/mps> (accessed on 10 August 2020).

¹³ The survey in question can be found here: <https://openstate.eu/nl/2019/08/verstoppertje-in-den-haag-tweede-kamerleden-vaak-onbereikbaar/> (accessed 10 August 2020). On the one hand, social desirability may bias self-reported responsiveness upwards. On the other hand, less than a third of all MPs responded in the survey and it is likely that the respondents are among the most responsive MPs, given that the survey itself was circulated via email. In the absence of any further information, we assume that these biases roughly cancel each other out.

¹⁴ <https://www.abgeordnetennetwatch.de/bundestag/abgeordnete> (accessed on 10 August 2020).

response rates reported on this website produce a remarkable intraclass correlation of 0.66, though this is likely to be inflated compared to responsiveness to private emails.¹⁵ On the basis of these numbers, we set the expected intraclass correlation to 0.15 for the United Kingdom – as a high response rate may well indicate low within-person variance – and to 0.30 for the other three countries.

Using these parameters, Table 3 presents the estimated level of statistical power for our main effects as a function of effect size and number of waves. Each cell is based on a thousand simulations. For example, if we assume a main effect size of five percentage points, we have a probability of 0.77 of detecting this effect with two waves in all countries. Given that this is close to the conventional level of statistical power (0.80), and given that we want to limit the time and effort that MPs spend answering our emails (for ethical reasons), we choose two waves.

Table 3. Statistical power for *main effects* as a function of effect size and number of waves

	Effect size (in percentage points)		
	2.5	5	10
One wave	0.22	0.53	0.97
Two waves	0.32	0.77	>0.99
Three waves	0.44	0.90	>0.99
Four waves	0.52	0.96	>0.99

Note: Numbers represent estimated power using $\alpha = 0.05$ (one-tailed).

In Table 4, we present so-called type-M errors, which show the ratio by which we overestimate the causal effect if this effect is statistically significant (Gelman and Carlin, 2014). This shows that type-M errors are very modest with two waves and an effect size of five percentage points; on average, a significant causal effect exaggerates the true effect by a factor of 1.15.

¹⁵ On the one hand, many German MPs answer all questions posed by citizens on the website. On the other hand, a substantial number of them do not participate in the website at all and have not answered a single question. Since public questions seem to induce a quite different response than private correspondence, we have not used this website in determining our expected response rate for Germany.

Table 4. Type-M errors for *main effects* as a function of effect size and number of waves

	Effect size (in percentage points)		
	2.5	5	10
One wave	2.62	1.49	1.02
Two waves	1.93	1.15	1.00
Three waves	1.61	1.07	1.00
Four waves	1.43	1.02	1.00

Note: Numbers represent estimated power using $\alpha = 0.05$ (one-tailed).

Tables 5 and 6 illustrate the issues with statistical power for the interaction effects focusing on the interaction between MPs' personal characteristics and treatment factors. Here, we have used non-academic education as a proxy for working-class backgrounds. Table 5 shows that an interaction between lower education among MPs and working-class citizens have to be very large (>10 percentage points) for us to detect this in a reliable way. This is even starker in Table 6, due the greater scarcity of minority ethnic MPs in our four countries. This demonstrates why we are cautious about this type of analysis and why we approach it in exploratory ways.

Table 5. Statistical power for *interaction effects between treatment factor and non-academic education* as a function of effect size and number of waves

	Effect size (in percentage points)			
	2.5	5	10	15
One wave	0.09	0.15	0.36	0.63
Two waves	0.13	0.26	0.60	0.88
Three waves	0.16	0.35	0.75	0.97
Four waves	0.18	0.44	0.86	0.99

Note: Numbers represent estimated power using $\alpha = 0.05$ (one-tailed).

Table 6. Statistical power for *interaction effects between treatment factor and minority ethnicity* as a function of effect size and number of waves

	Effect size (in percentage points)			
	2.5	5	10	15
One wave	0.05	0.09	0.20	0.41
Two waves	0.07	0.13	0.32	0.63
Three waves	0.08	0.19	0.42	0.74
Four waves	0.10	0.23	0.54	0.85

Note: Numbers represent estimated power using $\alpha = 0.05$ (one-tailed).

We conclude by noting that Tables 5 and 6 may present conservative estimates. As we argued above, effect sizes are likely to be larger with a more continuous dependent variable. Moreover, we have not used block assignment in these simulations that should improve the precision of our estimates. On the other hand, we have not excluded MPs who may be excluded from the actual data set because they request a postal address or follow-up conversation, as we do not know how common this will be. The bottom line is that, in view of these issues, we remain cautious and tap into context effects in exploratory ways.

11. Other design issues

The type of CSFE that we outlined in this pre-analysis plan raises ethical issues that have been subject to debate (e.g. Desposato, 2016). These concerns particularly result from the lack of informed consent among our experimental subjects. We move forward with this design since it inflicts minimal harm to legislators and thus does not raise moral issues similar to biomedical research. We also subjected our experiment to local reviews in the different countries in which we plan to run it to ensure that our design is able to maximize social benefits while minimizing costs. Since we actively intervene into the political process with this experiment, these costs particularly might concern the political community, i.e. voters. But they also might concern other researchers that could be subject to professional backlashes if MPs feel tricked or deceived.

In line with ethical guidelines, we do not send out emails to MPs beyond the two treatment waves. This is to say that we do not respond to MPs if they pose questions, make requests or even question the authenticity of our emails. The only exception to this occurs when a large group of MPs (for instance, in different parties) raises doubts about the authenticity of the emails. In this case, it may be necessary to terminate the experiment and send out a standardized explanation of our data collection to all MPs. Regardless of such contingencies, we will not analyze the data until the end of the response window of the second wave.

References

- Bansak, K. *et al.* (2019) ‘Beyond the Breaking Point? Survey Satisficing in Conjoint Experiments’, *Political Science Research and Methods*, online first.
- Bol, D. *et al.* (2020) ‘The Importance of Personal Vote Intentions for the Responsiveness of

- Legislators: A Field Experiment', *European Journal of Political Research*, online first.
- Boynton, G. R., Patterson, S. C. and Hedlund, R. D. (1969) 'The Missing Links in Legislative Politics: Attentive Constituents', *Journal of Politics*, 31(3), pp. 700–721.
- Broockman, D. E. (2013) 'Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives', *American Journal of Political Science*, 57(3), pp. 521–536.
- Butler, D. M. (2014) *Representing the Advantaged: How Politicians Reinforce Inequality*. New York: Cambridge University Press.
- Butler, D. M. and Broockman, D. E. (2011) 'Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators', *American Journal of Political Science*, 55(3), pp. 463–477.
- Butler, D. M. and Crabtree, C. (2017) 'Moving Beyond Measurement: Adapting Audit Studies to Test Bias-Reducing Interventions', *Journal of Experimental Political Science*, 4(1), pp. 57–67.
- Butler, D. M. and Dynes, A. M. (2016) 'How Politicians Discount the Opinions of Constituents with Whom They Disagree', *American Journal of Political Science*, 60(4), pp. 975–989.
- Butler, D. M., Karpowitz, C. F. and Pope, J. C. (2012) 'A Field Experiment on Legislators' Home Styles: Service versus Policy', *Journal of Politics*, 74(2), pp. 474–486.
- Cain, B. E., Ferejohn, J. A. and Fiorina, M. P. (1987) *The Personal Vote: Constituency Service and Electoral Independence*. Cambridge: Harvard University Press.
- Carey, J. M. (2007) 'Competing Principals, Political Institutions, and Party Unity in Legislative Voting', *American Journal of Political Science*, 51(1), pp. 92–107.
- Carnes, N. (2013) *White-Collar Government: The Hidden Role of Class in Economic Policy Making*. Chicago: University of Chicago Press.
- Carnes, N. and Holbein, J. (2019) 'Do Public Officials Exhibit Social Class Biases When They Handle Casework? Evidence from Multiple Correspondence Experiments', *PLoS ONE*, 14(3), pp. 1–9.
- Castles, F. G. and Wildenmann, R. (eds) (1986) *Visions and Realities of Party Government*. Berlin: De Gruyter.
- Coppock, A. (2019) 'Avoiding Post-Treatment Bias in Audit Experiments', *Journal of Experimental Political Science*, 6(1), pp. 1–4.
- Costa, M. (2017) 'How Responsive are Political Elites? A Meta-Analysis of Experiments on Public Officials', *Journal of Experimental Political Science*, 4(3), pp. 241–254.

- Cox, G. W. (1990) 'Centripetal and Centrifugal Incentives in Electoral Systems', *American Journal of Political Science*, 34(4), pp. 903-935.
- Dahl, R. A. and Tufte, E. R. (1973) *Size and Democracy*. Stanford: Stanford University Press.
- Dalton, R. J. (1985) 'Political Parties and Political Representation: Party Supporters and Party Elites in Nine Nations', *Comparative Political Studies*, 18(3), pp. 267-299.
- Desposato, S. (ed.) (2016) *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*. New York: Routledge.
- Erikson, R. and Goldthorpe, J. H. (1992) *The Constant Flux: A Study of Class Mobility in Industrial Societies*. London: Clarendon Press.
- Esaiasson, P., Gilljam, M. and Persson, M. (2013) 'Communicative Responsiveness and Other Central Concepts in Between-Election Democracy', in Esaiasson, P. and Narud, H. M. (eds) *Between-Election Democracy: The Representative Relationship After Election Day*. Colchester: ECPR Press, pp. 15-33.
- Esaiasson, P., Gilljam, M. and Persson, M. (2017) 'Responsiveness Beyond Policy Satisfaction: Does It Matter to Citizens?', *Comparative Political Studies*, 50(6), pp. 739-765.
- Ezrow, L. *et al.* (2010) 'Mean Voter Representation and Partisan Constituency Representation: Do Parties Respond to the Mean Voter Position or to their Supporters?', *Party Politics*, 17(3), pp. 275-301.
- Ganzeboom, H. B. G., De Graaf, P. M. and Treiman, D. J. (1992) 'A Standard International Socio-Economic Index of Occupational Status', *Social Science Research*, 21(1), pp. 1-56.
- Gelman, A. (2018) *You Need 16 Times the Sample Size to Estimate an Interaction than to Estimate a Main Effect*. Available at: <https://statmodeling.stat.columbia.edu/2018/03/15/need-16-times-sample-size-estimate-interaction-estimate-main-effect/> (Accessed: 22 September 2020).
- Gelman, A. and Carlin, J. (2014) 'Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors', *Perspectives on Psychological Science*, 9(6), pp. 641-651.
- Gerber, A. S. and Green, D. P. (2012) *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton & Company.
- Giger, N., Rosset, J. and Bernauer, J. (2012) 'The Poor Political Representation of the Poor in a Comparative Perspective', *Representation*, 48(1), pp. 47-61.
- Gilens, M. and Page, B. I. (2014) 'Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens', *Perspectives on Politics*, 12(3), pp. 565-581.

- Grohs, S., Adam, C. and Knill, C. (2016) 'Are Some Citizens More Equal than Others? Evidence from a Field Experiment', *Public Administration Review*, 76(1), pp. 155–164.
- Habel, P. and Birch, S. (2019) 'A Field Experiment on the Effects of Ethnicity and Socioeconomic Status on the Quality of Representation', *Legislative Studies Quarterly*, 44(3), pp. 389–420.
- Hainmueller, J., Hopkins, D. J. and Yamamoto, T. (2014) 'Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments', *Political Analysis*, 22(1), pp. 1–30.
- Heitshusen, V., Young, G. and Wood, D. M. (2005) 'Electoral Context and MP Constituency Focus in Australia, Canada, Ireland, New Zealand, and the United Kingdom', *American Journal of Political Science*, 49(1), pp. 32–45.
- Henderson, J. and Brooks, J. (2016) 'Mediating the Electoral Connection: The Information Effects of Voter Signals on Legislative Behavior', *Journal of Politics*, 78(3), pp. 653–669.
- Katz, R. S. (1986) 'Party Government: A Rationalistic Conception', in Castles, F. G. and Wildenmann, R. (eds) *Visions and Realities of Party Government*. Berlin: De Gruyter.
- Key, V. O. (1961) *Public Opinion and American Democracy*. New York: Alfred A. Knopf.
- Koehn, P. (2005) *European Parliament Proceedings Parallel Corpus 1996-2011*. Available at: <https://www.statmt.org/europarl/> (Accessed: 5 October 2020).
- Magni, G. and Ponce de Leon, Z. (2020) 'Women Want an Answer! Field Experiments on Elected Officials and Gender Bias', *Journal of Experimental Political Science*, online first.
- Manin, B. (1997) *The Principles of Representative Government*. New York: Cambridge University Press.
- Mansbridge, J. (1999) 'Should Blacks Represent Blacks and Women Represent Women? A Contingent "Yes"', *Journal of Politics*, 61(3), pp. 628–657.
- Mansbridge, J. (2003) 'Rethinking Representation', *American Political Science Review*, 97(4), pp. 515–528.
- Martin, P. S. (2003) 'Voting's Rewards: Voter Turnout, Attentive Publics, and Congressional Allocation of Federal Money', *American Journal of Political Science*, 47(1), pp. 110–127.
- Mayhew, D. R. (1974) *Congress: The Electoral Connection*. New Haven: Yale University Press.
- McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001) 'Birds of a Feather: Homophily in

- Social Networks', *Annual Review of Sociology*, 27, pp. 415–444.
- Miler, K. C. (2010) *Constituency Representation in Congress: The View from Capitol Hill*. Cambridge: Cambridge University Press.
- Muralidharan, K., Romero, M. and Wüthrich, K. (2020) *Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments*. National Bureau of Economic Research Working Paper No. 26562.
- Norris, P. (2004) *Electoral Engineering: Voting Rules and Political Behavior*. Cambridge: Cambridge University Press.
- Oesch, D. (2006) *Redrawing the Class Map: Stratification and Institutions in Britain, Germany, Sweden and Switzerland*. Basingstoke: Palgrave Macmillan.
- Pitkin, H. F. (1967) *The Concept of Representation*. Berkeley: University of California Press.
- Romeijn, J. (2018) 'Do Political Parties Listen to The(ir) Public? Public Opinion-Party Linkage on Specific Policy Issues', *Party Politics*, 26(4), pp. 426–436.
- Schakel, W., Burgoon, B. and Hakhverdian, A. (2020) 'Real but Unequal Representation in Welfare State Reform', *Politics & Society*, 48(1), pp. 131–163.
- Schneider, B. *et al.* (1998) 'Personality and Organizations: A Test of the Homogeneity of Personality Hypothesis', *Journal of Applied Psychology*, 83(3), pp. 462–470.
- Strøm, K. (1997) 'Rules, Reasons, and Routines: Legislative Roles in Parliamentary Democracies', in Müller, W. C. and Saalfeld, T. (eds) *Members of Parliament in Western Europe: Roles and Behavior*. London: Frank Cass.
- Wlezien, C. and Soroka, S. N. (2012) 'Political Institutions and the Opinion-Policy Link', *West European Politics*, 35(6), pp. 1407–1432.

Appendix A: Email registration protocol

We have registered the first available email address with Gmail for each name, using the following protocol (with the name Jessica Smith as an example).

1. jessicasmith@gmail.com
2. jessica.smith@gmail.com
3. jsmith@gmail.com
4. j.smith@gmail.com
5. jessicasmith@gmail.com (after this, we use other letters as middle initials, in the following order: m, a, j, s, r, l, c, d, g, k, h, e, b, t, n, p, f, w, i, v, o. These correspond to the most common middle initials between 1950 and 1990.)
6. jessica.m.smith@gmail.com (after this, we use other letters as middle initials, as under 5)
7. jmsmith@gmail.com (after this, we use other letters as middle initials, as under 5)
8. j.m.smith@gmail.com (after this, we use other letters as middle initials, as under 5)
9. If none of these ($4 + 4 * 21 =$) 88 accounts are available, we add “1” at the end of the name and start again at the top of this list (jessicasmith1@gmail.com, jessica_smith1@gmail.com, etc.).
10. Next, we add “2” – “9” at the end of the name and start again at the top of the list.
11. After this, a random three-digit number is added and we start from the top again.

Appendix B: Coding the quality of responses

Aim

As a dimension of responsiveness, we wish to measure the quality of the responses we receive. Quality is difficult to operationalize as it might depend on the nature of the request as well as the subjective perception of the receiver. In the following, we present a common framework for coding the quality of the responses to e-mail requests about policy matters.

Conceptual reference

The content of the responses provided by MPs are coded according to a communicative understanding of responsiveness (Esaïasson, Gilljam and Persson, 2013). This concept is a sensible reference, since it perceives responsiveness as contingent upon what legislators signal in their interactions with citizens. This is what we are interested in.

Esaïasson and co-authors (2017) suggest three forms of responsiveness in elite – mass communication:

1. To listen: actions taken to stay informed about citizen sentiments
2. To explain: actions taken to provide a credible justification for policy decisions
3. To adapt: actions taken to adjust policy decisions in the direction of the majority opinion

Conceptual frame for coding mail content

We adapt our conceptual reference to take the specific empirical context of our experiment into account. This includes adding a fourth form of responsiveness: position taking. This is important because we ask specifically about positions in our experiment and we expect MPs to respond to this request. Our adaptations also include changing the term ‘adaption’ to ‘action’. Responsiveness may involve more than just taking measures to adapt or adjust to voter preferences. Most importantly, the senders of our mails do not state any position the MP could adapt to. Our four empirical dimensions are:

1. *Listening* = signals of paying attention, seeking information and/or understanding the content of the request
2. *Position-taking* = clarification of policy position with regard to the relevant issue
3. *Action* = identification of a specific action that has been or will be taken in relation to the issue in general or to realize or promote stated position

4. *Explaining* = justification of stated policy position or action clarifying reasons or considerations for taking a specific position or action

A response of highest quality will thus include an indication of the MP having understood the nature of the request, offering a clear position with regard to the issue of the request, describing which actions will be taken to promote this position, and justifying any positions or actions by outlining substantial reasons for it.

Coding rules

To code the four dimensions, we ask coders to code explicit utterances in a categorical way, i.e. as being either true or false. We expect such simple dichotomous codes to increase reliability, which is crucial for our comparative study.

1. Listening

The act of listening is denoted by an explicit expression that a) references the content of the e-mail received, i.e. Covid crisis and challenges raised by it, and/or b) invites the sender to speak up, raise further concerns, and use specific channels of communication.

Please code if the response entails:

1 = Any expression of listening

0 = No expression of listening

For example, a response should be coded '1' if it includes statements like "Thank you for your e-mail regarding your concerns on covid-19", "Thank you for raising your concerns regarding the current corona crisis", "I understand your concerns regarding covid-19", "could you please elaborate on your concerns regarding the corona crisis?", "please visit my office hour to share your concerns about covid-19".

In contrast, responses not including such statements or only including statements like "thank you for your e-mail" or "would you please contact me to provide more information" without any indication of the MP engaging with the content of the email should be coded '0'.

2. Position taking

The act of position taking is denoted by any explicit and specific attitude or preference with regard to the management of the covid-19 situation and challenges raised by it.

Please code if the response entails:

1 = Any expression of position taking

0 = No expression of position taking

For example, a response should be coded '1' if it includes statements like: "I agree with the government policy, we need to keep strict regulation of social interaction", "We in the Conservative party, want to make sure that our businesses also survive corona, we want to limit restrictions on businesses as much as possible", or "The most important thing, is to keep the most vulnerable people in our society safe, we need to maintain regulation regarding our social interaction to do so."

In contrast, responses not including such statements or only including statements like "I support government policy", "This is very difficult, we have no clear solution", "I think we need a long term plan" without clarifying any policy position should be coded '0'.

3. Action

Action is denoted by any explicit reference to or description of specific individual or party level actions taken in relation to the management of covid-19. This could be formal parliamentary actions: ask parliamentary question, propose legislation etc., informal parliamentary action: contact minister, organize meeting among opposition parties etc., or extra-parliamentary action: contact interest groups, organize protests etc.

Please code if the response entails:

1 = Any description of action

0 = No description of action

For example, a response should be coded '1' if it includes a statement like: "I have negotiated all night with the minister to make him offer more help to our suffering artists", "I have asked 54 questions to the minister on when he will offer a long term plan", "We have arranged an informal hearing among all unions to understand how we support employment best" which all refer to specific actions.

In contrast, a response should be coded '0' if it includes no such reference or only very general references to action like "I work hard every day to handle the crisis" or "we do anything we can to get us out of the crisis".

4. Explaining

The act of explaining is denoted by any explicit justification for why specific positions on Covid-19 or specific actions are taken, i.e. providing reasons for why restrictions should be maintained, loosened or tightened with regard to Covid-19. Reasons could include e.g. concerns for elderly people, the economy or democratic rights.

Please code if the response entails:

1 = Any expression of explaining

0 = No expression of explaining

For example, two of the three position statements above also include a justification and therefore should be coded with “1” on this dimension: “We in the Conservative party, want to make sure that our businesses also survive corona, we want to limit restrictions on businesses as much as possible”, “The most important thing, is to keep the most vulnerable people in our society safe, we need to maintain regulation regarding our social interaction to do so.” Whereas the third position statement does not include a justification and thus should be coded with “0” on this dimension: “I agree with the government policy, we need to keep strict regulation of social interaction” doesn’t.

Generally, any response stating a substantial reason for a given position or action should be coded ‘1’ whereas responses without any justification for actions or positions should be coded ‘0’. Note that justifications need to be policy-specific, i.e. specific to the policy position taken. We do not code non-substantial “reasons” such as ‘because I support the government’, ‘because I respect my party position’, ‘because I do not have the sufficient expertise to question the current policy’.

Measurement

To measure the quality of the response we assume that each dimension is equally important and construct an additive index (response-quality) of the four empirical dimensions ranging from 0 to 4. This response-quality index will be dichotomized by the median and used for our overall index of responsiveness. Observations that are on the median will be recoded to either 0 or 1, depending on what results in the most even distribution. For example, if 45% of all observations are below the median and 35% are above it, the median value is grouped with the top 35%, since a 55-45 split is more even than a 65-35 split.

Appendix C: Translated treatment texts

Danish, version 1

Email subject: Vedrørende corona

Kære [Title] [MP's surname],

Mit navn er [first name] [surname]. Jeg arbejder som [rengøringsassistent / advokat], og jeg er bekymret for konsekvenserne af covid-19 krisen. Jeg arbejder i en stor virksomhed, og min egen situation er for øjeblikket tryk. Men jeg er bekymret for fremtiden. Mennesker omkring mig mister deres job eller må gå ned i løn. Samtidig bliver mange andre problemer overset, fordi alt handler om corona.

[Jeg støtter [party name] og / statement left out for control group] jeg vil gerne vide, hvad du og [party name] gør for at få os gennem denne krise på bedst mulig vis.

Jeg ser frem til at høre fra dig.

Med venlig hilsen,

[Constituent's name]

Email subject: Spørgsmål om corona

Kære [Title] [MP's surname],

Jeg er [first name] [surname]. Jeg skriver til dig, fordi jeg er bekymret for konsekvenserne af corona-krisen. Jeg er [rengøringsassistent / advokat], og jeg kender mange, som oplever vanskeligheder. De mister deres job eller må acceptere mindre i løn. Jeg er fortsat sikker, da jeg arbejder for en stor virksomhed. Men jeg er urolig, ikke kun for corona som sådan, men også fordi mange andre problemer ikke får den nødvendige opmærksomhed på grund af corona.

Jeg vil derfor gerne spørge dig, [som en der støtter [party name] / statement left out for control group] hvordan du og [party name] planlægger at styre os gennem disse vanskelige tider.

Jeg glæder mig til at høre dit svar.

Med venlig hilsen,

[Constituent's name]

Email subject: Vraag over de coronacrisis

Geachte [Title] [MP's surname],

Mijn naam is [first name] [surname]. Ik werk als [schoonmaker / advocaat] en ik maak me zorgen over de gevolgen van de coronacrisis. Ik werk bij een groot bedrijf en ik denk dat mijn eigen baan voorlopig nog veilig is. Maar ik ben wel bezorgd over de langere termijn. Ik ken veel mensen die hun baan hebben verloren of salaris moeten inleveren als gevolg van de crisis. En veel andere problemen krijgen nu weinig aandacht omdat alles over corona gaat.

[Als aanhanger van [party name], zou ik / Ik zou] graag willen weten wat u en [party name] gaan doen om ons zo goed mogelijk uit deze crisis te krijgen.

Ik kijk uit naar uw reactie.

Met vriendelijke groet,

[Constituent's name]

Email subject: Coronavirus

Geachte [Title] [MP's surname],

Ik ben [first name] [surname]. Ik e-mail u omdat ik bezorgd ben over de impact van de coronacrisis. Ik ben [schoonmaker / advocaat] en ik zie veel mensen om me heen die zijn getroffen door de gevolgen van deze crisis, omdat ze hun baan kwijtraken of nu minder verdienen. Omdat ik voor een groot bedrijf werk voel ik me voorlopig nog veilig. Maar ik maak me zorgen over de toekomst. Dat komt niet alleen door corona, maar ook omdat veel andere problemen nu weinig aandacht krijgen.

Ik wil u vragen [, als aanhanger van [party name], / statement left out for control group] hoe u en [party name] van plan zijn om ons door deze moeilijke periode te leiden.

Ik ben benieuwd naar uw antwoord.

Vriendelijke groeten,

[Constituent's name]

Email subject: Corona

Sehr geehrte(r) [Title] [MP's surname],

Ich heiße [first name] [surname] und bin erst in diesem Jahr in Ihren Wahlkreis gezogen. Ich bin als [Reinigungskraft / Anwalt(in)] tätig und bin besorgt über die Folgen der Corona Pandemie. Ich arbeite für ein größeres Unternehmen und bin persönlich noch nicht betroffen. Aber ich mache mir Sorgen über die längerfristigen Folgen. Viele Menschen in meinem Umfeld sind schon betroffen, indem sie ihren Arbeitsplatz verloren haben oder mit weniger Einkommen leben müssen. Außerdem gibt es viele andere Probleme, die wir vernachlässigen, weil sich alles um Corona dreht.

[Als Anhänger der [party name] wüsste ich / Ich wüsste] gerne, was Sie und die [party name] unternehmen, um uns durch diese Krise zu bringen.

Ich freue mich auf Ihre Antwort.

Mit freundlichen Grüßen,

[Constituent's name]

Email subject: Frage zu Coron

Sehr geehrte(r) [Title] [MP's surname],

Mein Name ist [first name] [surname]. Im letzten Jahr bin ich in Ihren Wahlkreis gezogen. Ich schreibe Ihnen, weil ich mir Sorgen über die Folgen der Corona Krise mache. Ich bin als [Reinigungskraft / Anwält(in)] tätig und sehe viele in meinem Umfeld, die ihren Arbeitsplatz verloren haben oder über weniger Einkommen verfügen. Ich arbeite in einem größeren Unternehmen und bin im Moment nicht betroffen. Aber ich mache mir Sorgen über die Zukunft. Ich bin auch unsicher wegen der vielen anderen Probleme, die durch Corona weniger Aufmerksamkeit bekommen.

[Ich bin Anhänger der [party name]. Deshalb will ich / Ich will] Sie fragen, was Sie und die [party name] unternehmen, um uns durch diese schwierigen Zeiten zu bringen.

Ich freue mich auf Ihre Antwort.

Mit freundlichen Grüßen,

[Constituent's name]

