Online Appendix for

“When is a pledge a pledge?”

Contents

[Appendix A: Literature 2](#_Toc63545035)

[A1: Short discussion on existing literature 3](#_Toc63545036)

[Appendix B: Argument 5](#_Toc63545037)

[B1: Detailed discussion of theoretical dimensions 6](#_Toc63545038)

[Appendix C: Experimental design 12](#_Toc63545039)

[C1: Data collection 13](#_Toc63545040)

[C2: Survey questionnaires 14](#_Toc63545041)

[C2a: Survey questionnaire, United States 14](#_Toc63545042)

[C2b: Survey questionnaire, Britain 16](#_Toc63545043)

[C2c: Survey questionnaire, Denmark 18](#_Toc63545044)

[C3: Discussions on case selection 20](#_Toc63545045)

[C4: Detailed description of conjoint attributes 21](#_Toc63545046)

[Appendix D: Main analysis 25](#_Toc63545047)

[D1: Regression table of main results 26](#_Toc63545048)

[D2: All attributes in same model 27](#_Toc63545049)

[D3: Alternative model specifications 28](#_Toc63545050)

[Appendix D3a: Logit models 29](#_Toc63545051)

[Appendix D3b: Random effects models 30](#_Toc63545052)

[Appendix D3c: Respondent fixed effects models 31](#_Toc63545053)

[Appendix D3d: Respondent and round fixed effects models 32](#_Toc63545054)

[D4: Predicted probabilities of specific statements 33](#_Toc63545055)

[D4a: Specific statements (United States) 33](#_Toc63545056)

[D4b: Specific statements (Britain) 34](#_Toc63545057)

[D4c: Specific statements (Denmark) 35](#_Toc63545058)

[Appendix D5: Split-samples across gender, age group, education, and political views 36](#_Toc63545059)

[Appendix D5a: Main models with individual-level controls 37](#_Toc63545060)

[Appendix D5b: Split-sample analysis with females only 38](#_Toc63545061)

[Appendix D5c: Split-sample analysis with males only 39](#_Toc63545062)

[Appendix D5d: Split-sample analysis with respondents below 39 years only 40](#_Toc63545063)

[Appendix D5e: Split-sample analysis with respondents between 40-65 years only 41](#_Toc63545064)

[Appendix D5f: Split-sample analysis with respondents above 65 years only 42](#_Toc63545065)

[Appendix D5g: Split-sample analysis with low educational level respondents only 43](#_Toc63545066)

[Appendix D5h: Split-sample analysis with high educational level respondents only 44](#_Toc63545067)

[Appendix D5i: Split-sample analysis with left-leaning respondents only 45](#_Toc63545068)

[Appendix D5j: Split-sample analysis with right-leaning respondents only 46](#_Toc63545069)

[Appendix D6: Separate analyses across issues 47](#_Toc63545070)

[Appendix D6a: Main models only for “Health” statements 48](#_Toc63545071)

[Appendix D6b: Main models only for “Immigration” statements 49](#_Toc63545072)

[Appendix D6c: Main models only for “Labor” statements 50](#_Toc63545073)

[Appendix D6d: Main models only for “Olympics” statements 51](#_Toc63545074)

[Appendix D7: Examining the direct effect of senders 52](#_Toc63545075)

[Appendix D8: Examining ordering effects 53](#_Toc63545076)

[Appendix D8a: Main results only for Round 1 54](#_Toc63545077)

[Appendix D8b: Main results only for Round 2 55](#_Toc63545078)

[Appendix D8c: Main results only for Round 3 56](#_Toc63545079)

[Appendix D8d: Main results only for Round 4 57](#_Toc63545080)

[Appendix D9: Examining interaction between attributes through split-samples 58](#_Toc63545081)

[Appendix D9a: Main models only for “We promise” statements 59](#_Toc63545082)

[Appendix D9b: Main models only for “We think there should be” statements 60](#_Toc63545083)

[Appendix D9c: Main models only for no conditions statements 61](#_Toc63545084)

[Appendix D9d: Main models only for “if opportunity” statements 62](#_Toc63545085)

[Appendix D9e: Main models only for output statements 63](#_Toc63545086)

[Appendix E: Mediation analysis 64](#_Toc63545087)

[Appendix E1: Mediation analysis of conditionality attributes 66](#_Toc63545088)

[Appendix E2: Mediation analysis of all significant attributes 67](#_Toc63545089)

[Appendix F: Accountability analysis 68](#_Toc63545090)

[Appendix F1: Cross-correlation of election pledge and accountability 69](#_Toc63545091)

[Appendix F2: Controlling for party choice 70](#_Toc63545092)

[Appendix F2a: Controlling for sender 71](#_Toc63545093)

[Appendix F2b: Controlling for sender’s status as in government or opposition 72](#_Toc63545094)

[Appendix F2c: Split-sample, opposition party senders only 73](#_Toc63545095)

[Appendix F2d: Split-sample, government party senders only 74](#_Toc63545096)

[References 75](#_Toc63545097)

# Appendix A: Literature

This section provides a short discussion of existing literature.

## A1: Short discussion on existing literature

Systematic research into election pledges dates back several decades (e.g., Royed 1996; Thomson 2001; Mansergh and Thomson 2007; Thomson et al. 2017; Naurin, Royed and Thomson 2019). Until recently, the main focus was on the link between pledges and policies. The main takeaway from this line of work is that governments on average tend to keep their pledges, but that the pledge-keeping track record varies across countries. In Britain, where single-party majority governments are typically the rule, 80–85% of pledges are kept, whereas just 50–60% are kept in the United States (Thomson et al. 2017; see also Naurin 2009, 2014; Håkansson and Naurin 2016). Such high pledge-keeping records might seem surprising, which is likely a byproduct of newspapers’ dominant focus on broken, rather than fulfilled, election pledges (Müller 2020).

Another strand of research turns away from elite activities of pledge-making and pledge-keeping and instead studies what voters know about pledges and whether they use them to hold governments accountable. Thomson (2011), for instance, explores Irish voters’ knowledge of four pledges issued in the 2002 election campaign. The author documents comparably low levels of knowledge, though it should be noted that respondents were quizzed at the time of the 2007 election, five years after the pledges were made. Naurin, Soroka, and Markwat (2019), another prominent example, employ Swedish post-election survey experiments to gauge how citizens hold the government accountable for kept and broken pledges, respectively. They find a substantial negativity bias wherein broken pledges are given much more weight than kept pledges. Matthieß (2020) forcefully demonstrates that government parties are in fact more likely to be rewarded by voters come election time if they have kept more pledges during their previous term in office.

As attention in the research community pivots towards the voter level, a new question is coming into view: Which characteristics make voters categorize some statements as pledges and others not? Without exact information about voters’ definitions of a pledge, it is difficult to tell whether our current understanding of voters’ knowledge and use of pledges is a function of a mismatch between how scholars define pledges when developing their surveys, on the one hand, and voters’ definitions, on the other hand. More importantly, given the strong normative force of pledges, policymakers should know which of their statements voters categorize as pledges.

Examining this, Naurin’s (2011) qualitative work with in-depth interviews of Swedish citizens shows—among other things—that these interviewees tended to understand pledges quite broadly (e.g., as doing what is best for society). Dupont et al. (2019) add to Naurin (2011) and show that the degree of commitment in a statement affects the likelihood of it being considered a pledge. Both studies demonstrate that voters distinguish between political statements and that they are able to reason about how they do so. Moreover, recent linguistic work on the subject demonstrates that voters’ understanding of pledges—much like other forms of political communication—is affected by the use of certain value-laden words (Lindgren 2017, 2018; Lindgren and Naurin 2017).

# Appendix B: Argument

This section presents a detailed discussion of our argument. The main points are the same as in the main article, but we discuss each dimension in more detail and engage more deeply with relevant literature.

## B1: Detailed discussion of theoretical dimensions

As summarized in Table 1 in the main text, we conceptualize the perceptual logic of election pledges as structured along four rhetorical dimensions—commitment, the pledge giver, policy content, and quantification. These four dimensions cover all rhetorical aspects of what scholars normally define as pledges, as can be seen when we relate the four dimensions to the conventional definition of a pledge as “a statement committing [i.e. *commitment*] a party [i.e. *the pledge giver*] to one specific action or outcome [i.e. *policy content*] that can be clearly determined to have occurred or not [i.e. *quantification*]” (Thomson et. al. 2017: 532).

In the context of electoral politics, we can see the pledge as a credible commitment device that allows parties to trade votes today for policies in the future. Voters may want to vote prospectively, i.e., based on desired future policy, but such prospective voting comes with a significant risk. The vote is cast on Election Day, but the return only arrives afterwards. How can the voter be sure that a party, should it win, will not pursue some other, unwanted policy? The solution is the pledge.

Pledges work as credible commitment devices for two reasons. First, breaking them makes the pledge giver less trustworthy. Since trustworthiness is important in electoral politics (Funk 1996; 1999), the willingness to jeopardize it signals sincere commitment. Second, pledges are public, which means that it is not for the pledge giver to decide whether a pledge has been kept or not (Schedler 1998; Royed et al. 2019). In short, pledges signal sincere commitment because the stakes are high, and the yardstick of fulfillment is, at least partially, outside the control of the pledge giver.

Because of this essential feature of pledges, the most basic dimension of an election pledge to voters is the degree of *commitment* (Schedler 1998: 198-199; Royed et al. 2019). A pledge that does not contain an explicit commitment, such as “we promise,” “we vow,” or “we swear,” is in everyday language less likely to be labeled a pledge at all, and we assume that voters share this intuition when it comes to election pledge perceptions. Much the same can be said when a pledge giver attaches conditions to the pledge. By saying that a pledge will only be fulfilled under certain conditions, such as “if the economy allows it” or “if the opportunity arises,” the pledge giver is essentially creating back doors to escape the credible commitment of the pledge. Obviously, some conditions are very reasonable and may be regarded so by voters, but the strength of commitment, all else equal, declines when conditions are attached. The degree of commitment is vital because it signals the extent to which a pledge giver sincerely prioritizes the issue at stake. Governments are invariantly confronted with numerous demands on their attention and always have to relentlessly prioritize which problems to focus on (Jones and Baumgartner 2005). In the everyday tussle of post-election politics, a pledge is a way of attaching special importance to a specific problem. Conversely, if voters do not believe that a pledge giver honestly prioritizes an issue, even a strong rhetorical commitment is unlikely to make voters consider a statement to be a pledge. This leads to the first two hypotheses about the direct and mediated effects of strong rhetorical commitment:

**Hypothesis 1direct:**Commitment increases voters’ perception that a statement is a pledge.

**Hypothesis 1mediated:** The effect of commitment is mediated by voters’ perception that the pledge giver priorities the issue.

The second dimension of pledges is the *pledge giver*. Terminologically, the identity of the person who makes a statement should not affect whether voters consider the statement a pledge to begin with. In this sense, pledge perceptions are neutral. Yet political psychology tells us that voters often harbor a partisan bias, meaning that they tend to interpret information in ways that align their partisan biases and incoming information so as to avoid so-called cognitive dissonance (Lodge and Taber 2013; Bisgaard 2019). We may speculate that this is true for pledges as well. On the one hand, voters may consider statements from their own party to be pledges simply because they trust that the politician will act according to her statement. On the other hand, since the trustworthiness of the pledge giver is at stake, partisans may be less willing to categorize a statement from their own party as a pledge because it creates a possible tension between the voter’s positive evaluation of the pledge giver today and the risk of having to change that evaluation tomorrow. This implies that the expected effect of the identity of the pledge giver can be both positive and negative:

**Hypothesis 2direct:**Voters’ perception that a statement is a pledge depends on the identity of the pledge giver.

**Hypothesis 2mediated:**The effect of the identity of the pledge giver is mediated by voters’ perception that the pledge giver is trustworthy.

The third dimension is the *policy* *content* of the pledge. Parties can make promises about virtually anything, but a crucial dividing line exists between output and outcome (Schedler 1998: 196-197; Naurin 2011: 127-128; Royed et al. 2019). Output is the direct consequence of a legislative decision; outcome is the desired end result of this decision. Politics is normally motivated by outcomes, such as a more industrious society, better public health, preservation of the national culture, or improved living standards for the jobless. Yet outcomes cannot be manipulated directly by the government. The government may hope that more spending or revised regulation will create the desired outcome, but the chances that it will can be quite low. Appreciating that some phenomena are less under government control than others, we can reasonably expect that voters differentiate between outputs and outcomes and that they do so by perceiving statements on outputs as pledges to a greater extent than statements on outcomes.

**Hypothesis 3direct:**Voters are more likely to consider a statement a pledge if the policy content concerns outputs rather than outcomes.

**Hypothesis 3mediated:** The effect of policy content is mediated by voters’ perception that the pledge giver has control over the issue.

The fourth and final dimension is *quantification*, i.e., whether the statement includes a numeric goal (e.g., not only promising more spending but 25% more spending) or time horizon (e.g., within the first year of the election) that makes it possible to unambiguously observe after the election whether a pledge has been kept (Schedler 1998: 196; Royed et al. 2019: 30). By proposing an exact yardstick for pledge-keeping before the election, parties in principle renounce the right to define their own pledge-keeping record. Instead, objective and specific metrics increase voters’ own ability to assess a pledge, i.e., the voters’ own sense of internal efficacy increases. By this logic, voters perceive quantified statements to be election pledges because they allow them to assess the pledge-keeping record. This is not to say that quantifying statements are the only way to make statements testable or that quantification always increases voters’ ability to test the fulfillment of pledges. Yet, in general, we should expect quantifiable messages to increase voters’ sense of efficacy—all else equal—compared to statements that are not specified numerically.

**Hypothesis 4direct:**Voters are more likely to consider a statement a pledge if the statement is

quantified.

**Hypothesis 4mediated:**The effect of quantification is mediated by voters’ perception of efficacy.

The Perceptual Logic of Pledges

|  |  |  |  |
| --- | --- | --- | --- |
| **Rhetorical dimension** | **Voters’ perception of…** | **Election pledge** | **Accountability** |
| Commitment | Priority | Is the statement a pledge? | Is sender accountable? |
| Pledge giver | Trustworthiness |
| Policy content | Control |
| Quantification | Efficacy |

So far, the discussion has focused on how the four rhetorical dimensions can affect voters’ perceptions of priority, trustworthiness, control, and efficacy, and thereby whether a statement is considered a pledge, as summarized in the table above (see also Table 1 in the main text). Commitment signals that the pledge giver prioritizes the issue sincerely; the pledge giver’s identity or affiliation determines the perceived trustworthiness of the statement; the policy content may be more or less under the control of the pledge giver; and quantification is relevant as it induces voters to believe that they have the ability to assess the statement.

Yet, theoretically, pledges are significant because voters imbue them with an expectation of fulfillment, which they do not do with statements that are not regarded as pledges. In other words, when a statement is considered a pledge, the actor making the statement should also be considered accountable for seeing it implemented. This particular connection between pledges and accountability is so far unexplored in the literature. What is more, based on our discussion of the perceptual logic of pledges, we may not only ask whether there is an association between categorizing a statement as a pledge, on the one hand, and believing a sender is responsible for implementing it, on the other hand; we may also ask which particular perceptions are important for establishing this association.

To a large extent, similar logics may prevail. We may speculate that a sender is considered accountable if the actor is using strong committing language rather than a loose, non-committing formulation. The fit between the voters’ party identification and the sender may similarly affect the degree to which a sender is accountable, as parties or candidates a voter identifies with are seen as less accountable. Finally, the policy content and quantification may also affect the extent to which a sender is considered accountable since voters are presumably less forgiving regarding pledges on issues that are easier to control and more visible than those that are not.

**Hypothesis 5:**Commitment, the identity of the pledge giver, policy content, and quantification

have similar effects on voters’ perceptions that a statement is a pledge and that the pledge giver should be held accountable for the statement.

# Appendix C: Experimental design

This section presents the data collection process and full questionnaires in the United States, Britain, and Denmark, and discuss the reasons for choosing the three countries. In addition, it presents the conjoint attributes in more detail, and assesses ordering effects.

## C1: Data collection

|  |  |  |  |
| --- | --- | --- | --- |
|  | **United States** | **Britain** | **Denmark** |
| **Data collection period** | Start: 18-02-2019  End: 25-02-2019 | Start: 12-12-2018  End: 19-12-2018 | Start: 12-12-2018  End: 19-12-2018 |
| **Total number of respondents** | 2032 | 2214 | 2031 |
| **Survey company** | YouGov | YouGov | YouGov |

## C2: Survey questionnaires

### C2a: Survey questionnaire, United States

**[q1] In general, how satisfied are you with the way democracy works in the US?**

* 0 – Very dissatisfied
* 1
* …
* 9
* 10 – Very satisfied
* Don’t know

**[q2] How much do you feel you can trust the following?**

The Republicans

The Democrats

Donald Trump

Nancy Pelosi

* 1 – Not at all
* 2
* 3
* 4
* 5 – Very much
* Don’t know

**[q3] How much do you agree or disagree with the following statements?**

Congress should work to secure a higher standard of living for the unemployed

Congress should work to secure better service in the health care system

Congress should work to secure a stricter immigration policy

Congress should work to secure better performances by American athletes

* Disagree completely
* Disagree partially
* Neither agree nor disagree
* Agree partially
* Agree completely
* Don’t know

**[q4] How important is each of the following topics to you?**

The living conditions for the unemployed

The health care system

Immigration

The performance of American athletes

* Not at all important
* Not important
* Neither/nor
* Important
* Very important
* Don’t know

|  |
| --- |
| *Vignette example:*  In connection with the latest midterm election, the Republicans made the following statement: “We promise 25% more cancer screenings within the next year if the economy allows it.” |

**[q5] Would you consider this an election pledge?**

* Yes
* No
* Don’t know

**[q6] How much do you agree with the following statement?**

This is a high priority for [sender].

* Disagree completely
* Disagree
* Neither/nor
* Agree
* Agree completely
* Don’t know

**[q7] Is this something [sender] can do something about if [sender] really wants to?**

* Definitely not
* Probably not
* Neither/nor
* Probably
* Definitely
* Don’t know

**[q8] Should [sender] be held accountable for whether this is implemented?**

* Yes
* No
* Don’t know

**[q9] How much do you agree with the following statement?**

In reality, it is impossible for people like me to check whether [sender] actually implements this.

* Disagree completely
* Disagree partially
* Neither agree nor disagree
* Agree partially
* Agree completely
* Don’t know

**[q10] We only have one last question. Which of the following topics have you not read about?**

* Agriculture
* Cancer
* The Olympics
* Unemployed people

### C2b: Survey questionnaire, Britain

**[q1] In general, how satisfied are you with the way democracy works in England?**

* 0 – Very dissatisfied
* 1
* …
* 9
* 10 – Very satisfied
* Don’t know

**[q2] How much do you feel you can trust the following?**

The Conservative Party

The Labour Party

The UK Independence Party (UKIP)

The Liberal Democrats

Theresa May

Jeremy Corbyn

* 1 – Not at all
* 2
* 3
* 4
* 5 – Very much
* Don’t know

**[q3] How much do you agree or disagree with the following statements?**

Congress should work to secure a higher standard of living for the unemployed

Congress should work to secure better service in the health care system

Congress should work to secure a stricter immigration policy

Congress should work to secure better performances by English athletes

* Disagree completely
* Disagree partially
* Neither agree nor disagree
* Agree partially
* Agree completely
* Don’t know

**[q4] How important is each of the following topics to you?**

The living conditions for the unemployed

The health care system

Immigration

The performance of English athletes

* Not at all important
* Not important
* Neither/nor
* Important
* Very important
* Don’t know

|  |
| --- |
| *Vignette example:*  In connection with the latest parliamentary election, the Conservative Party made the following statement: “We promise 25% more cancer screenings within the next year if the economy allows it.” |

**[q5] Would you consider this an election pledge?**

* Yes
* No
* Don’t know

**[q6] How much do you agree with the following statement?**

This is a high priority for [sender].

* Disagree completely
* Disagree
* Neither/nor
* Agree
* Agree completely
* Don’t know

**[q7] Is this something [sender] can do something about if [sender] really wants to?**

* Definitely not
* Probably not
* Neither/nor
* Probably
* Definitely
* Don’t know

**[q8] Should [sender] be held accountable for whether this is implemented?**

* Yes
* No
* Don’t know

**[q9] How much do you agree with the following statement?**

In reality, it is impossible for people like me to check whether [sender] actually implements this.

* Disagree completely
* Disagree partially
* Neither agree nor disagree
* Agree partially
* Agree completely
* Don’t know

**[q10] We only have one last question. Which of the following topics have you not read about?**

* Agriculture
* Cancer
* The Olympics
* Unemployed people

### C2c: Survey questionnaire, Denmark

**[q1] Hvor tilfreds er du generelt med den måde, demokratiet fungerer på i Danmark?**

* 0 – Særdeles utilfreds
* 1
* …
* 9
* 10 – Særdeles tilfreds
* Ved ikke

**[q2] Hvor stor tillid mener du, man kan have til følgende?**

Partiet Venstre

Partiet Socialdemokratiet

Dansk Folkeparti

Radikale Venstre

Lars Løkke Rasmussen

Mette Frederiksen

* 1 – Overhovedet ingen tillid
* 2
* 3
* 4
* 5 – Meget stor tillid
* Ved ikke

**[q3] Hvor enig eller uenig er du i følgende udsagn?**

Folketinget bør sikre en højere levestandard for de arbejdsløse

Folketinget bør sikre et højere serviceniveau i sundhedsvæsenet

Folketinget bør sikre en strammere indvandrerpolitik

Folketinget bør sikre bedre danske sportspræstationer

* Helt uenig
* Delvist uenig
* Hverken enig eller uenig
* Delvist enig
* Helt enig
* Ved ikke

**[q4] Hvor vigtige er følgende emner for dig?**

De arbejdsløses vilkår

Sundhedsvæsenet

Indvandringen

Danske sportspræstationer

* Slet ikke vigtigt
* Ikke vigtigt
* Hverken/eller
* Vigtigt
* Meget vigtigt
* Ved ikke

|  |
| --- |
| *Vignette example:*  I forbindelse med seneste folketingsvalg udtalte Venstre: “Vi vil love 25% flere kræft-scanninger indenfor det næste år, hvis økonomien tillader det.” |

**[q5] Vil du mene, at dette er et valgløfte?**

* Ja
* Nej
* Ved ikke

**[q6] Hvor enig er du i følgende udsagn:**

Dette er noget [sender] prioriterer meget højt?

* Helt uenig
* Delvist uenig
* Hverken enig eller uenig
* Delvist enig
* Helt enig
* Ved ikke

**[q7] Er dette noget [sender] kan gøre noget ved, hvis [sender] virkeligt ønsker det?**

* Helt sikkert ikke
* Sandsynligvis ikke
* Hverken/eller
* Sandsynligvis
* Helt sikkert
* Ved ikke

**[q8] Bør [sender] holdes ansvarlig for, om dette bliver gennemført?**

* Ja
* Nej
* Ved ikke

**[q9] Hvor enig er du i følgende udsagn:**

Det er reelt umuligt for folk som mig, at kontrollere om [sender] faktisk gennemfører dette.

* Helt uenig
* Delvist uenig
* Hverken enig eller uenig
* Delvist enig
* Helt enig
* Ved ikke

**[q10] Vi har blot ét sidste spørgsmål. Hvilket af disse emner har du ikke læst om?**

* Landbrug
* Kræft
* OL
* Arbejdsløse

## C3: Discussions on case selection

In assessing the attributes of a political statement that induce citizens to consider it an election pledge, we use a conjoint experimental design (Hainmueller, Hopkins and Yamamoto 2014) on a representative sample of around 2,000 respondents each in the United States, Britain, and Denmark. We used YouGov as our survey institute, which, through its online panel, offers nationally representative samples on gender, age, geography, education, vote choice, and political viewpoints (see Appendix C2 for detailed questionnaire).

We study these three countries because they are home to widely different political cultures and institutions (Lijphart 2012; Dalton 2019) and because, as explained above, they vary in terms of the likelihood that governments keep their pledges (Thomson et al. 2017; Naurin, Royed and Thomson 2019). The United States has witnessed a highly polarized political climate since at least the 1980s, which intensified with the elections of President Obama in 2008 and President Trump in 2016. Furthermore, the many institutional veto points and bottlenecks of the United States’ system make it difficult for incumbents to deliver what they promise. Meanwhile, although the political climate of Britain has also become polarized recently, not least with the Brexit debate, the institutional environment is very different from that of the United States. Typically, the government has a majority in Parliament, meaning that it is considerably easier to pass legislation. Finally, Denmark is known for its consensus democracy. Perhaps because governments seldom have a majority of their own, compromise-seeking is the norm. Compared to the United States, policy gridlock is therefore much less a problem. If our findings are similar across these three nations, we feel confident that the results speak to the experience of Western democracies more broadly.

## C4: Detailed description of conjoint attributes

In the conjoint experiments, we presented our respondents with a series of political statements that randomly varied on the key attributes discussed above. Based on this, the respondents declared whether they perceived the statement to be an election pledge or not. Table 2 in the article presents all attributes and attribute values of the statements that we vary in the experiment. In many cases, the particular choices of attribute values might seem a bit arbitrary. Yet we have gone to great lengths ensure that the values capture each attribute dimension to the greatest extent possible and that respondents understand each attribute as intended. Specifically, we tested a range of different formulations on colleagues from our department and modified the wordings in cases where people seemed to misunderstand the wording or found it to be unclear. Moreover, we ran pilot studies in Denmark and Britain respectively in order assess respondents’ understandings of the attributes before running our main data collection.

These efforts led to the following operations. We vary the degree of commitment in two ways. First, we manipulate the *committing statement* from explicit promises to vague opinion declarations: “We promise,” “We will fight for,” “We will work towards,” and “We think there should be.” Based on our previous work on election manifestos in European democracies, we selected these four formulations as they are both a) commonly used by politicians and b) general enough to be used across our three countries. Also importantly, there is a clear hierarchy in the degree of commitment among the four declarations.

Second, we vary the *conditionality* of a statement by ending it with either no attached condition or a condition such as “if the economy allows it,” “if a majority in Congress is in favor” (“Parliament” in Britain and Denmark), or “if the opportunity arises.” Again, these three formulations are commonly used and respondents understand them perfectly in all three countries. In contrast to the committing statements, we do not expect a clear hierarchy between these three formulations, but we do expect that attaching a condition matters versus not attaching one.

We also manipulate the pledge giver. In each country, the *sender* could be either the governing party, the opposition party, or the political leader of either a center party (only Britain and Denmark) or a right-wing party (only Britain and Denmark). For example, in Britain, the sender is either “The Conservative Party,” “The Labour Party,” “Theresa May,” “Jeremy Corbyn,” “The Liberal Democrats,” or “United Kingdom Independence Party” (UKIP). In the main analysis below, we examined whether a given respondent votes for one of these senders, but we also ran additional models that examined whether the identity of the sender in itself has an effect. Including the main government party and leader and the main opposition party and leader make sense, as these are the most important and active players in the daily political process. We included small center parties in order to see whether statements from less dominant parties have similar impacts on people’s perceptions. Finally, given the debate about how populist parties seem able to get away with statements and actions that more mainstream parties could not, we include the right-wing Dansk Folkeparti in Denmark and the United Kingdom Independence Party (UKIP)in Britain.

Regarding the policy contentof the statement, we manipulated the proposed policy to take form as either an *output* or an *outcome*. Specifically, we distinguished between outputs such as “more cancer screenings,” “more resources to control family reunifications,” “higher unemployment benefits for the unemployed,” and “bigger budget for [American/British/Danish] Olympic athletes” and policy outcomes such as “more cancer survivors,” “fewer immigrants,” “more unemployed without economic problems,” and “more Olympic medals for [the United States/Britain/Denmark].” We deliberately included a diverse set of issues—immigration, unemployment protection, health care, and sports—to capture as much of the universe of real-life politics as possible.

We varied the quantification of the statement in two ways. First, we manipulated its *numeric goal* by attaching no numeric goal, a “10%,” a “25%,” or a “50%” qualification to the statements. Second, we also manipulated the *time horizon* in the political statement by distinguishing between proposed policies that will be implemented without any time constraints, “within the next year,” “within the next four years,” “within the next ten years,” or “in the future.” These numbers were chosen because they seem more realistic in combination with the chosen outputs and outcomes while still maintaining substantial numerical variation and a clear hierarchy among the different levels.

Finally, we varied the *timing* of the statement, specifying that it was made “in connection with the latest midterm election” (“parliamentary election” in Britain and Denmark), “last year,” or “last month.” Although not part of our analytical framework discussed above, we include this variable to increase the realism of the vignettes. Election pledges—while most often an election phenomenon—are also often made outside election campaigns. Yet given that this is not part of the main argument, we do not discuss the results of this attribute in the main article.

Complete Conjoint Design

|  |  |  |
| --- | --- | --- |
| **Dimension** | **Attributes** | **Attribute values (components)** |
|  | Timing | 1. In connection with the latest midterm/parliamentary election 2. Last year 3. Last month |
| *Pledge giver* | Sender | 1. the Republicans / the Conservative Party / the Liberal Party 2. the Democrats / the Labour Party / the Social Democrats 3. the Liberal Democrats / the Social Liberal Party 4. UKIP / the Danish People’s Party 5. Donald Trump / Theresa May / Lars Løkke Rasmussen 6. Nancy Pelosi / Jeremy Corbyn / Mette Frederiksen |
| *Commitment* | Committing statement | made the following statement:   1. “We promise 2. “We will fight for 3. “We will work towards 4. “We think there should be |
| *Quantification* | Numeric goal | 1. 10% 2. 20% 3. 50% |
| *Policy content* | Output/outcome | 1. more cancer survivors (outcome) 2. more cancer screenings (output) 3. fewer immigrants (outcome) 4. more resources to control family reunifications (output) 5. more unemployed without economic problems (outcome) 6. higher unemployment benefits for the unemployed (output) 7. more Olympic medals for the United States (outcome) 8. bigger budget for American Olympic athletes (output) |
| *Quantification* | Time horizon | 1. within the next year 2. within the next four years 3. within the next ten years 4. in the future |
| *Commitment* | Conditionality | 1. .” 2. if the economy allows it.” 3. if a majority in Congress / parliament is in favor.” 4. if the opportunity arises.” |

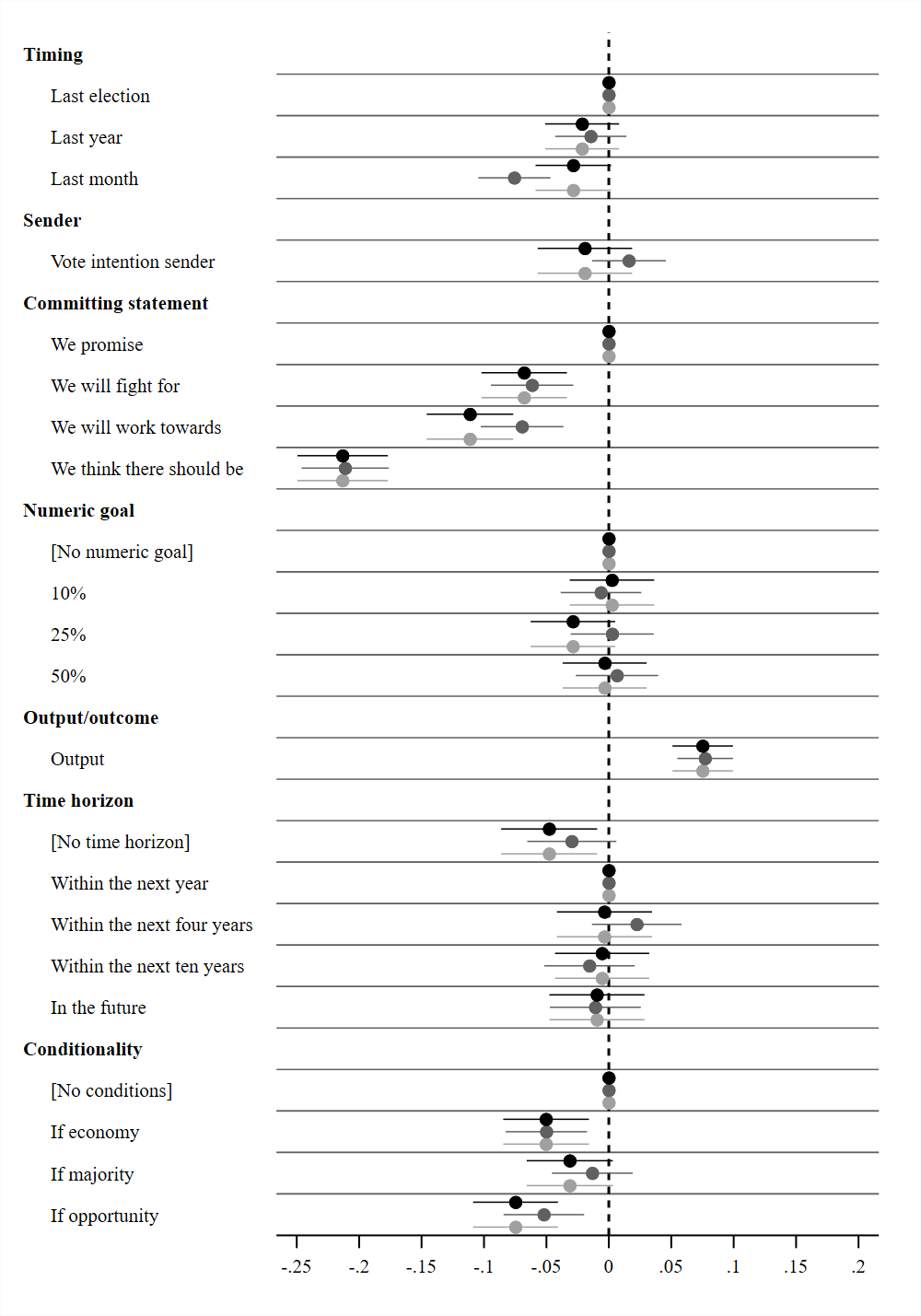
# Appendix D: Main analysis

This section presents an array of robustness checks by re-running the main analyses using alternative model specifications and estimation methods (note that these models present the full results of all attributes, including the timing dimension). In addition, it presents a range of extra results such as predicted probabilities for particular statements, and it presents several split-sample analyses that seek to assess whether results are similar across different experimental conditions or individual-level variables.

## D1: Regression table of main results

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | | |
|  | United States | | Britain | | Denmark | |
|  | (1) | | (2) | | (3) | |
| **Sender** |  |  | |  | |
| Vote sender | -0.017 | | 0.012 | | -0.016 | |
|  | (0.015) | | (0.015) | | (0.020) | |
|  |  | |  | |  | |
| **Committing statement** |  |  | |  | |
| Will fight for | -0.044\* | | -0.064\* | | -0.068\* | |
|  | (0.019) | | (0.017) | | (0.017) | |
|  |  | |  | |  | |
| Will work towards | -0.047\* | | -0.073\* | | -0.112\* | |
|  | (0.019) | | (0.017) | | (0.018) | |
|  |  | |  | |  | |
| Think there should be | -0.203\* | | -0.214\* | | -0.215\* | |
|  | (0.019) | | (0.018) | | (0.019) | |
|  |  | |  | |  | |
| **Numeric goals** |  |  | |  | |
| 10% | 0.002 | | -0.007 | | 0.005 | |
|  | (0.019) | | (0.017) | | (0.018) | |
|  |  | |  | |  | |
| 25% | -0.009 | | 0.007 | | -0.029 | |
|  | (0.019) | | (0.017) | | (0.017) | |
|  |  | |  | |  | |
| 50% | 0.031 | | 0.008 | | -0.005 | |
|  | (0.019) | | (0.017) | | (0.017) | |
|  |  | |  | |  | |
| **Output/outcome** |  |  | |  | |
| Output | 0.079\* | | 0.077\* | | 0.075\* | |
|  | (0.014) | | (0.012) | | (0.012) | |
|  |  | |  | |  | |
| **Time horizon** |  |  | |  | |
| [No time] | 0.017 | | -0.029 | | -0.056\* | |
|  | (0.021) | | (0.019) | | (0.020) | |
|  |  | |  | |  | |
| Next four years | 0.016 | | 0.021 | | -0.006 | |
|  | (0.021) | | (0.019) | | (0.020) | |
|  |  | |  | |  | |
| Next ten years | 0.006 | | -0.015 | | -0.009 | |
|  | (0.021) | | (0.019) | | (0.019) | |
|  |  | |  | |  | |
| In the future | -0.009 | | -0.011 | | -0.014 | |
|  | (0.021) | | (0.019) | | (0.020) | |
|  |  | |  | |  | |
| **Conditionality** |  |  | |  | |
| If economy | -0.040\* | | -0.048\* | | -0.054\* | |
|  | (0.019) | | (0.017) | | (0.018) | |
|  |  | |  | |  | |
| If majority | -0.002 | | -0.011 | | -0.030 | |
|  | (0.018) | | (0.017) | | (0.018) | |
|  |  | |  | |  | |
| If opportunity | -0.053\* | | -0.051\* | | -0.077\* | |
|  | (0.019) | | (0.017) | | (0.018) | |
|  |  | |  | |  | |
| Observations | 5611 | | 7017 | | 6296 | |

Standard errors (clustered on respondents) in parentheses \* *p* < 0.05. Reference categories: Last election (Timing), We promise (Committing statement), No numeric goal (Numeric goals), Within the next year (Time horizon), and No conditions (Conditionality). Constants are not reported in order to save space.

D2: All attributes in same model

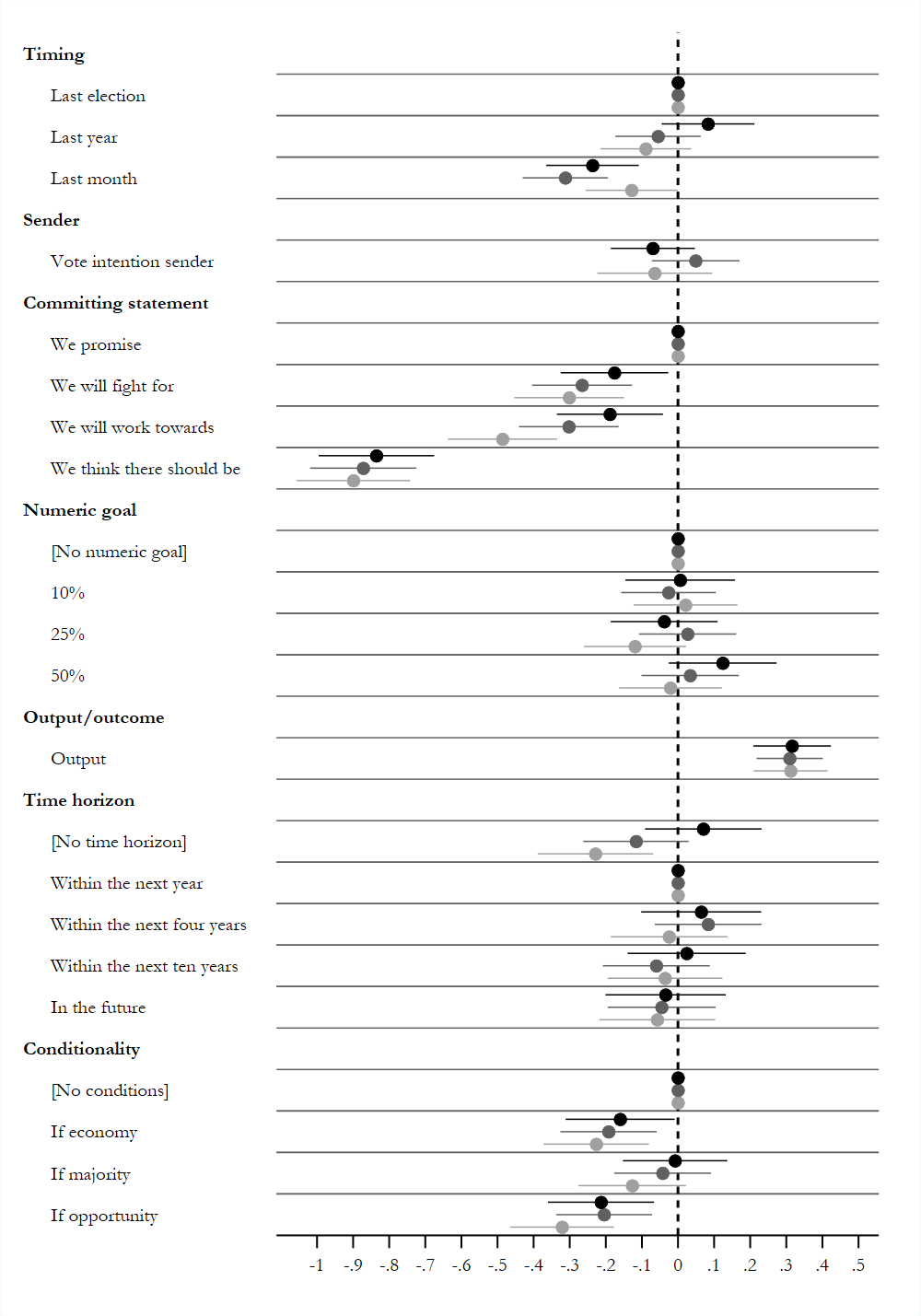
**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

## D3: Alternative model specifications

Despite the binary nature of the dependent variable, OLS models come with presentational advantages as the coefficients of each attribute can be easily compared across separate models. Still, in order to ensure that our conclusions are not simply an artifact of this estimation strategy, we present results from logit models as well. Moreover, we also present OLS models with respondent fixed effects that only account for within-respondent variation on treatment and outcome, i.e., whether a given respondent becomes more or less likely to consider a given political statement to be an election pledge as she encounters different political statements across the four rounds. These models control for all individual characteristics that are constant across all four rounds (which is basically all relevant pre-treatment individual characteristics). Finally, we present models that include both respondent fixed effects and round fixed effects. All results yield very similar findings.

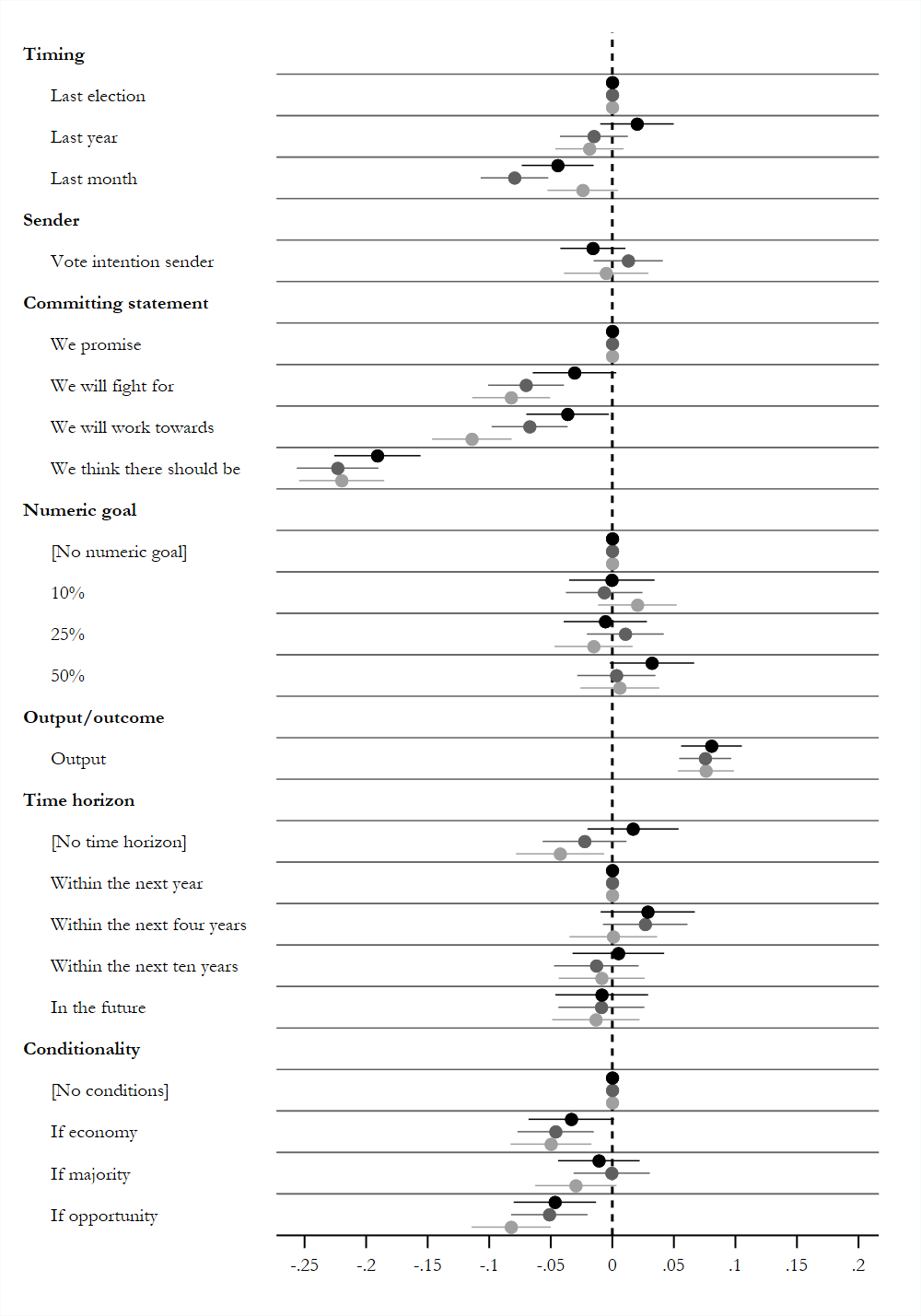
### Appendix D3a: Logit models



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

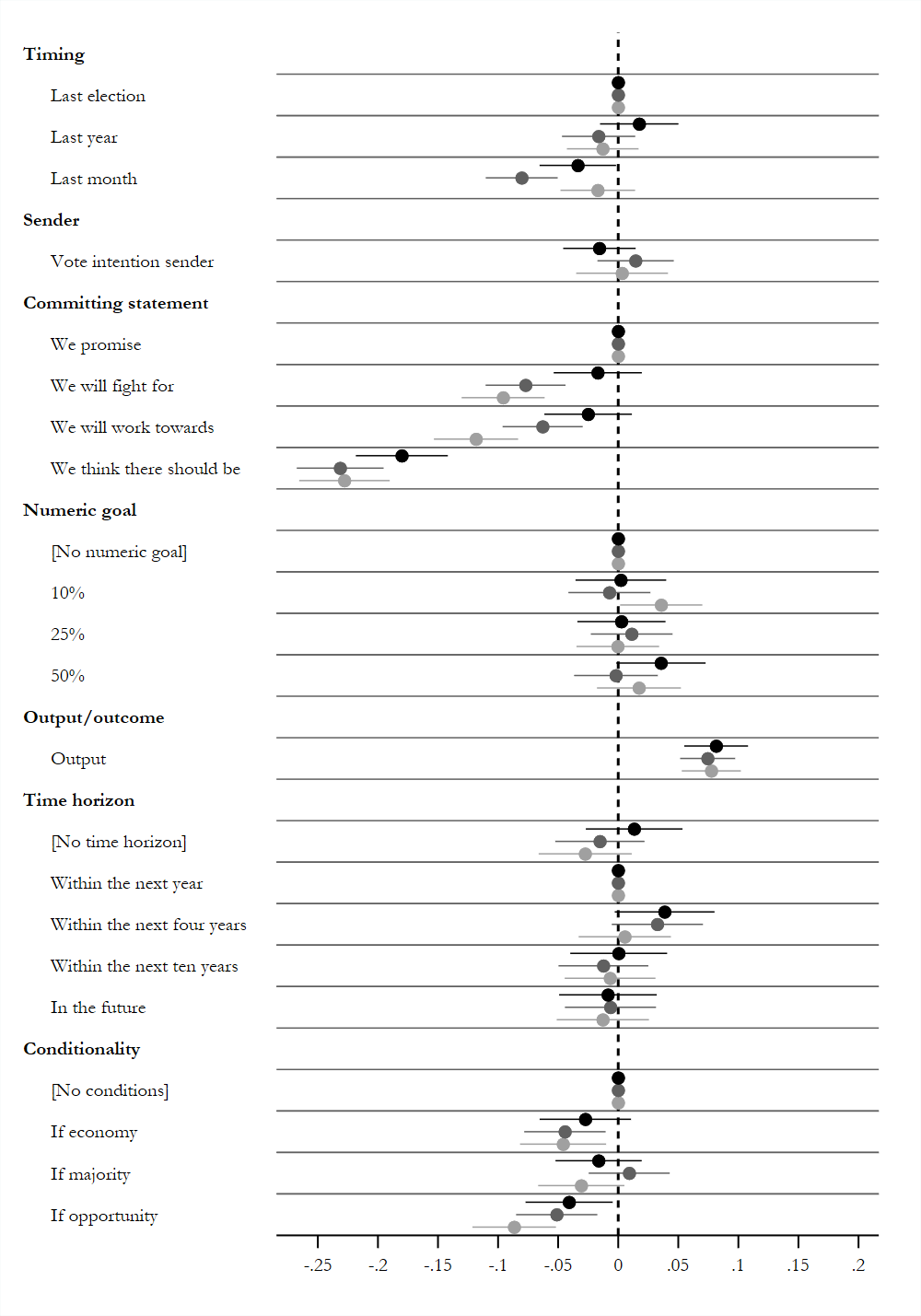
### Appendix D3b: Random effects models



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

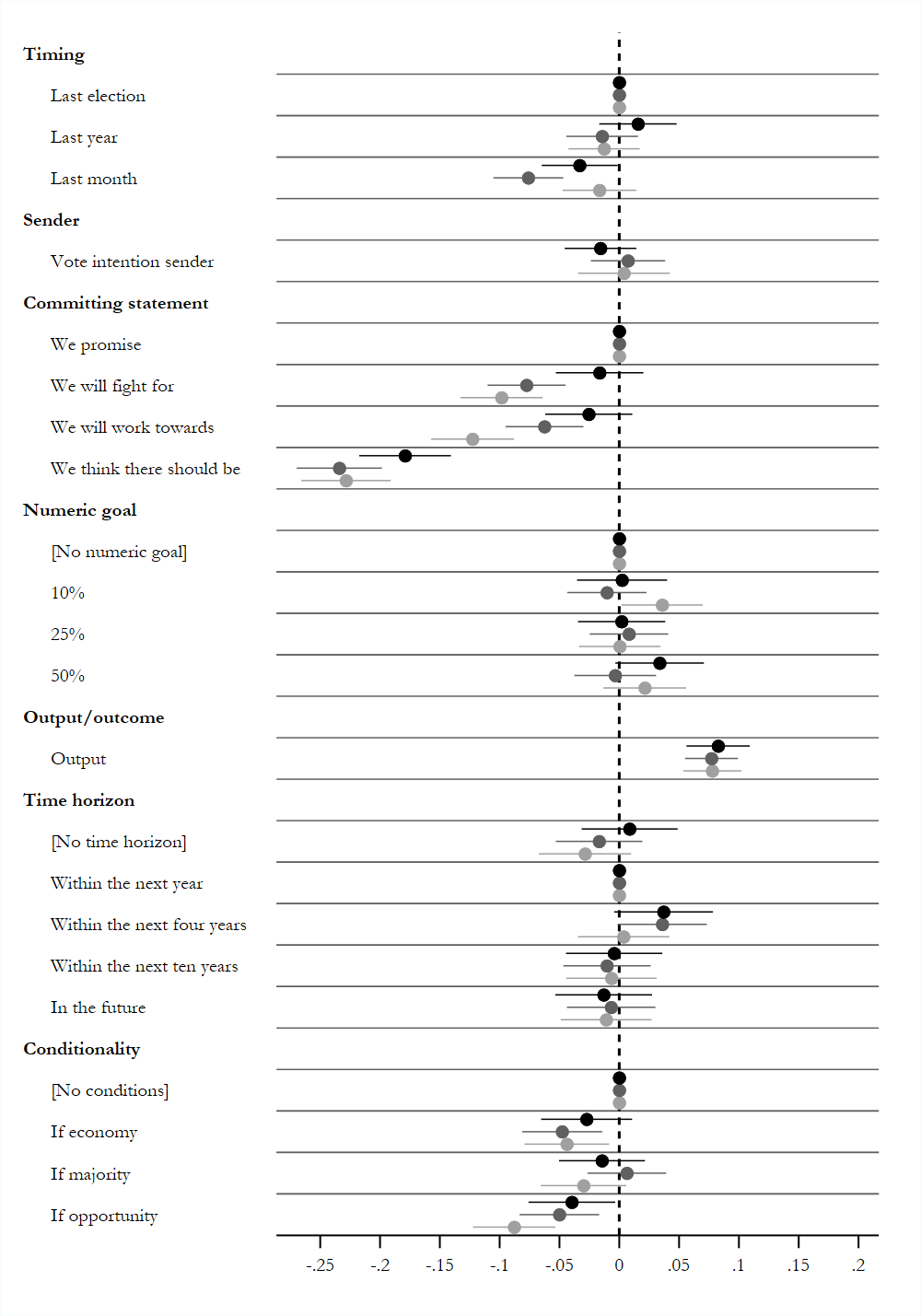
### Appendix D3c: Respondent fixed effects models



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

### Appendix D3d: Respondent and round fixed effects models

****

**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

## D4: Predicted probabilities of specific statements

### D4a: Specific statements (United States)

|  |  |  |
| --- | --- | --- |
| **Statement** | | **Probability** |
| **A** | Last year, the Republicans made the following statement: “We promise 10% more cancer screenings within the next four years if a majority in Congress is in favor.” | 61% |
| **B** | Last year, the Republicans made the following statement: “We will work towards 10% more cancer screenings within the next four years if a majority in Congress is in favor.” | 57% |
| **C** | Last year, the Republicans made the following statement: “We will work towards 10% more cancer screenings within the next four years if the economy allows it.” | 53% |
| **D** | Last year, the Republicans made the following statement: “We will work towards 10% more cancer survivors within the next four years if the economy allows it.” | 46% |
| **E** | Last year, the Republicans made the following statement: “We think there should be 10% more cancer survivors within the next four years.” | 35% |

**Note:** Predicted probabilities of a given statement being classified as an election pledge based on the American sample.

### D4b: Specific statements (Britain)

|  |  |  |
| --- | --- | --- |
| **Statement** | | **Probability** |
| **A** | Last year, the Conservative Party made the following statement: “We promise 10% more cancer screenings within the next four years if a majority in parliament is in favor.” | 73% |
| **B** | Last year, the Conservative Party made the following statement: “We will work towards 10% more cancer screenings within the next four years if a majority in parliament is in favor.” | 66% |
| **C** | Last year, the Conservative Party made the following statement: “We will work towards 10% more cancer screenings within the next four years if the economy allows it.” | 63% |
| **D** | Last year, the Conservative Party made the following statement: “We will work towards 10% more cancer survivors within the next four years if the economy allows it.” | 55% |
| **E** | Last year, the Conservative Party made the following statement: “We think there should be 10% more cancer survivors within the next four years.” | 46% |

**Note:** Predicted probabilities of a given statement being classified as an election pledge based on the British sample.

### D4c: Specific statements (Denmark)

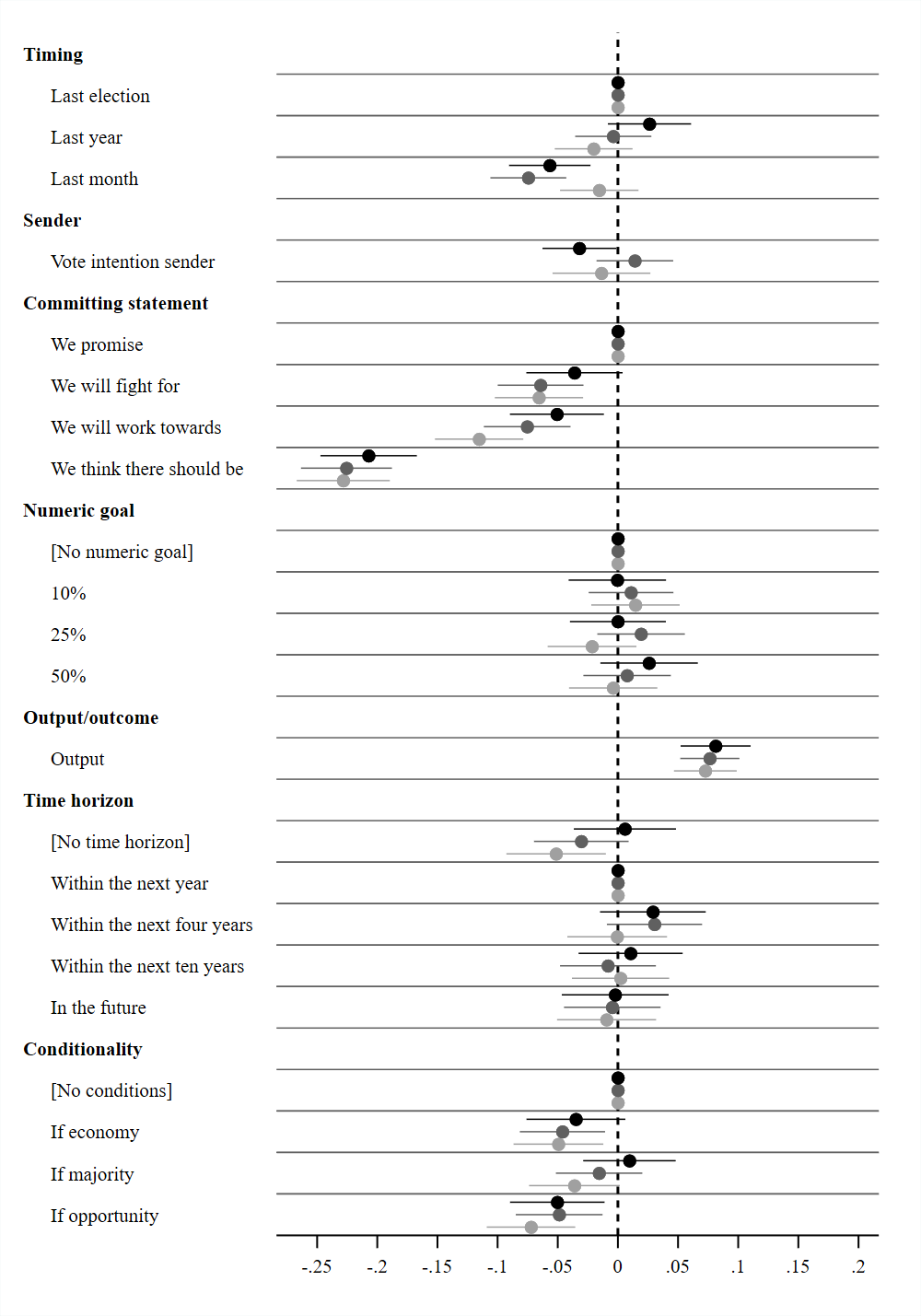
|  |  |  |
| --- | --- | --- |
| **Statement** | | **Probability** |
| **A** | Last year, the Liberal Party made the following statement: “We promise 10% more cancer screenings within the next four years if a majority in parliament is in favor.” | 76% |
| **B** | Last year, the Liberal Party made the following statement: “We will work towards 10% more cancer screenings within the next four years if a majority in parliament is in favor.” | 66% |
| **C** | Last year, the Liberal Party made the following statement: “We will work towards 10% more cancer screenings within the next four years if the economy allows it.” | 64% |
| **D** | Last year, the Liberal Party made the following statement: “We will work towards 10% more cancer survivors within the next four years if the economy allows it.” | 56% |
| **E** | Last year, the Liberal Party made the following statement: “We think there should be 10% more cancer survivors within the next four years.” | 50% |

**Note:** Predicted probabilities of a given statement being classified as an election pledge based on the Danish sample.

## Appendix D5: Split-samples across gender, age group, education, and political views

This section examines whether the results are similar across individual characteristics. We do so by undertaking a range of split-sample estimations for each individual category: gender, age, education, and political views. Although such estimations are generally quite inefficient, the results are remarkably consistent overall. We also provide results where all individual characteristics are simply controlled for.

### Appendix D5a: Main models with individual-level controls

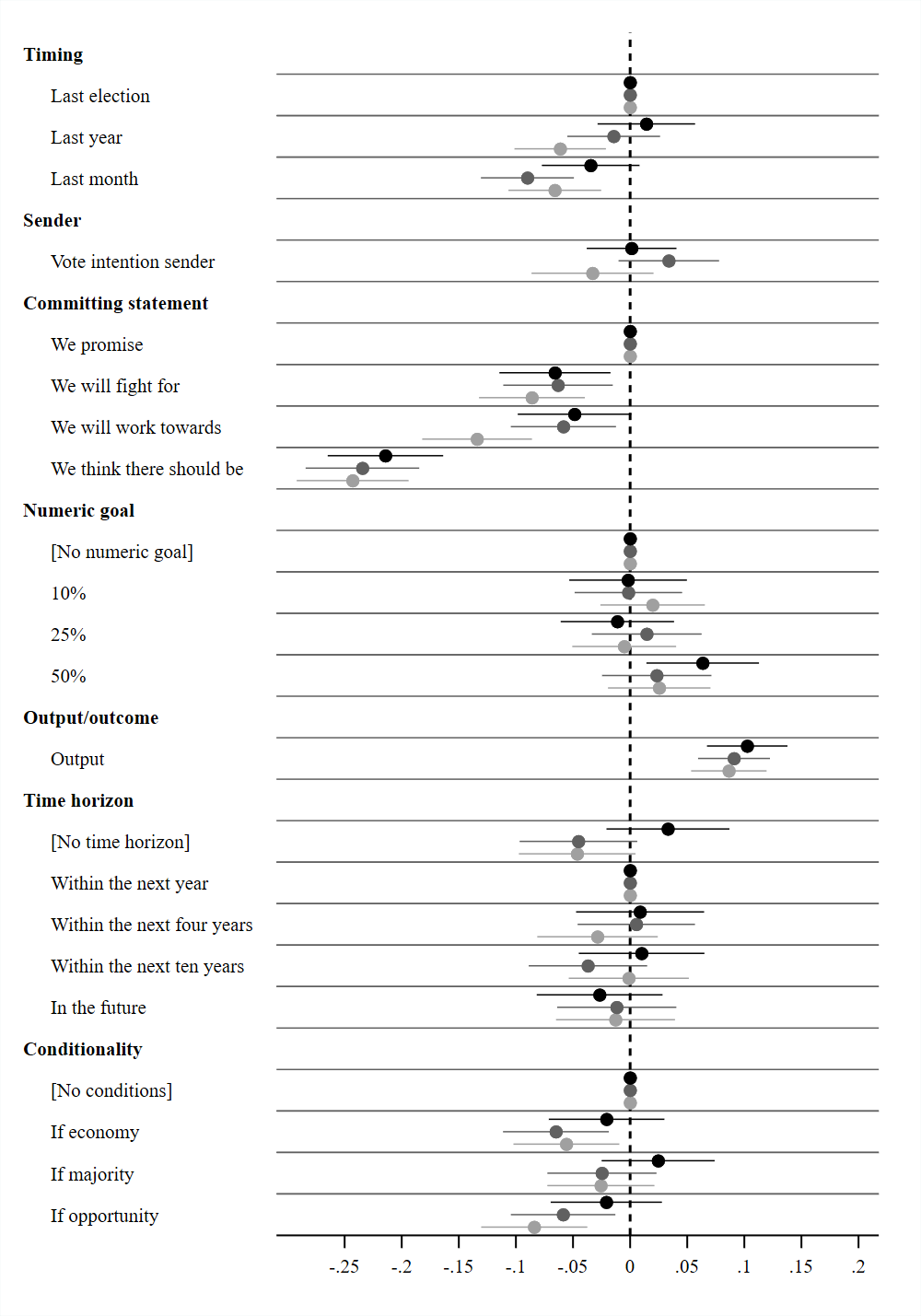


**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

Included controls: Gender, age, education, and self-placement on left-right scale.

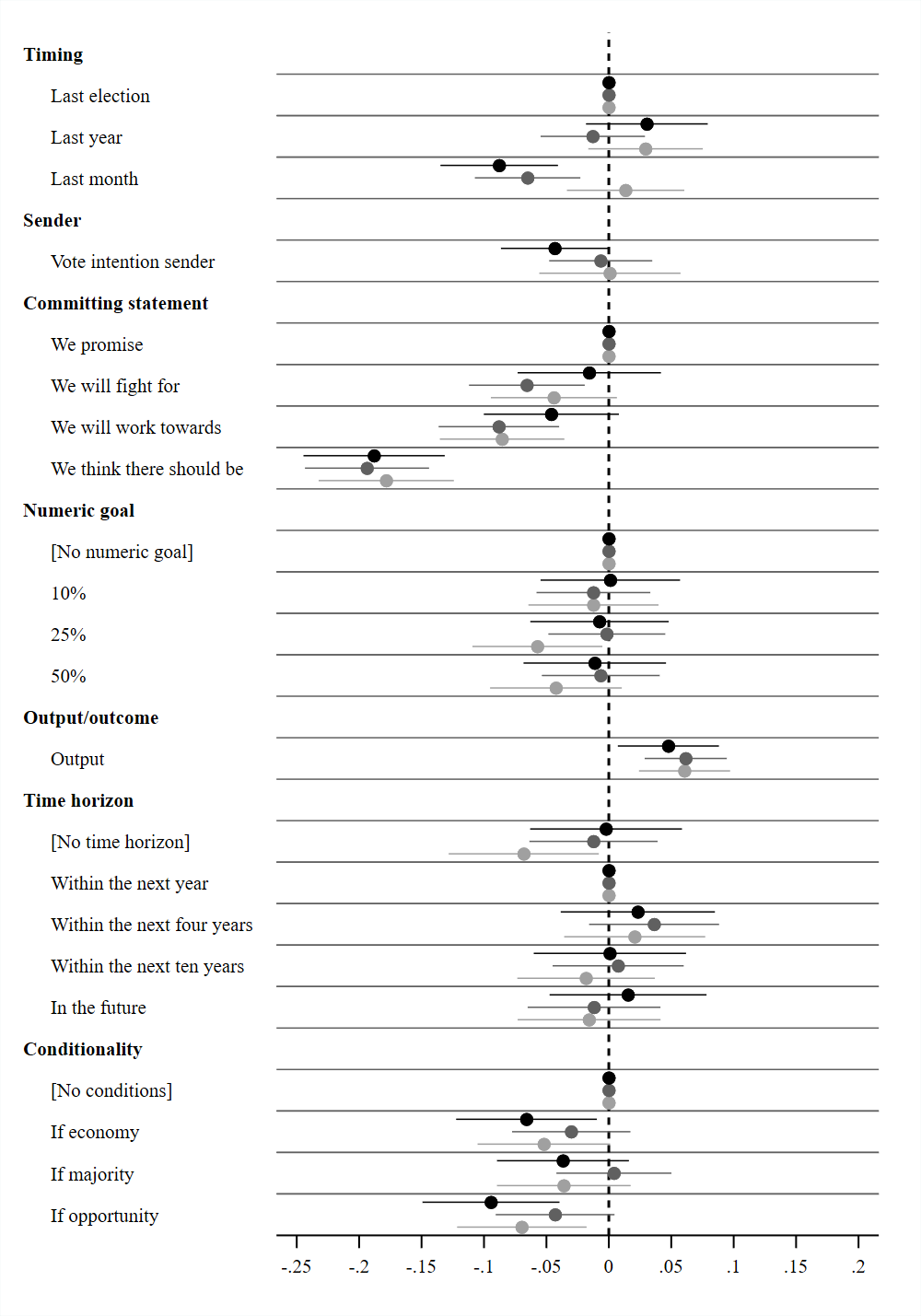
### Appendix D5b: Split-sample analysis with females only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

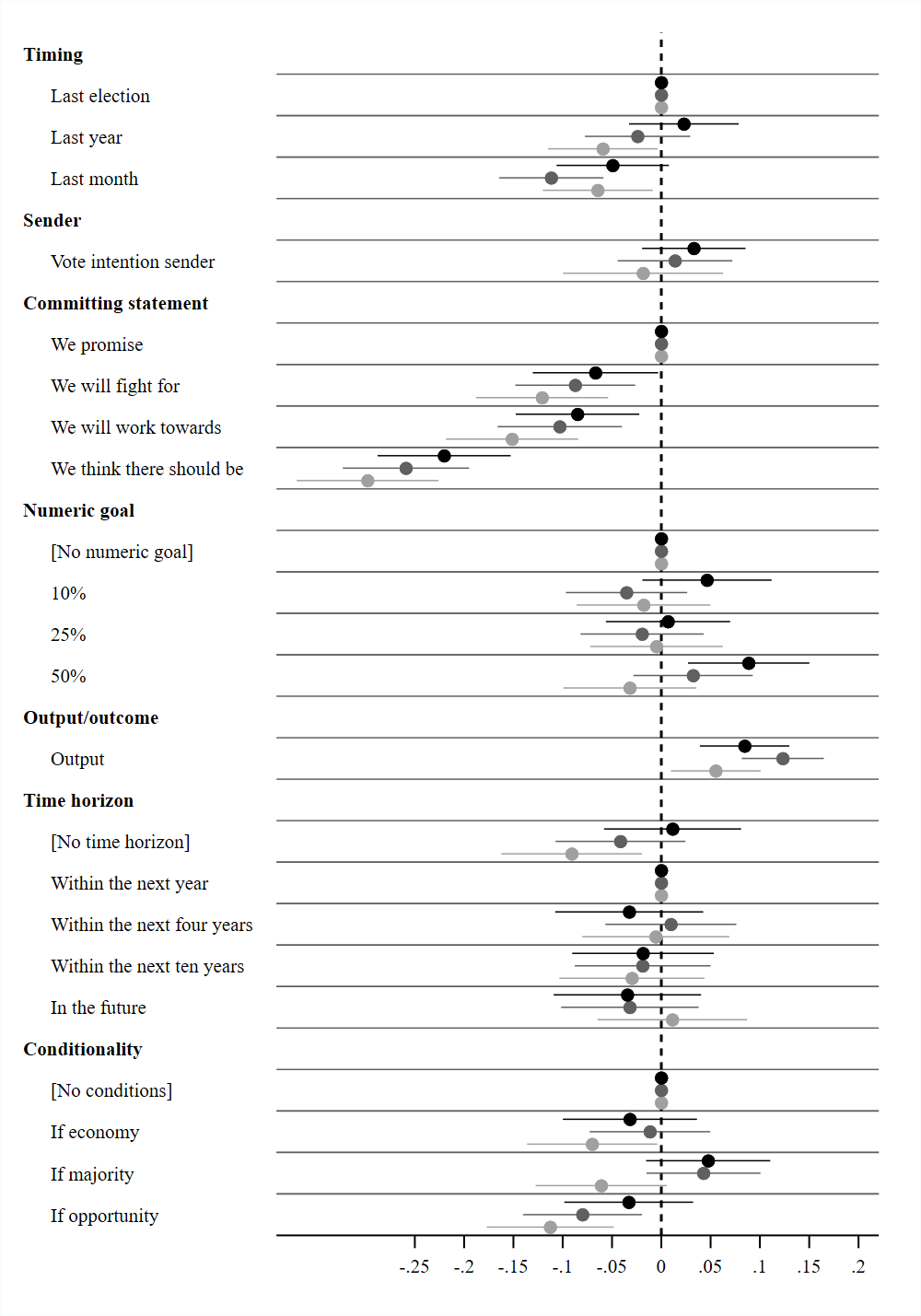
### Appendix D5c: Split-sample analysis with males only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

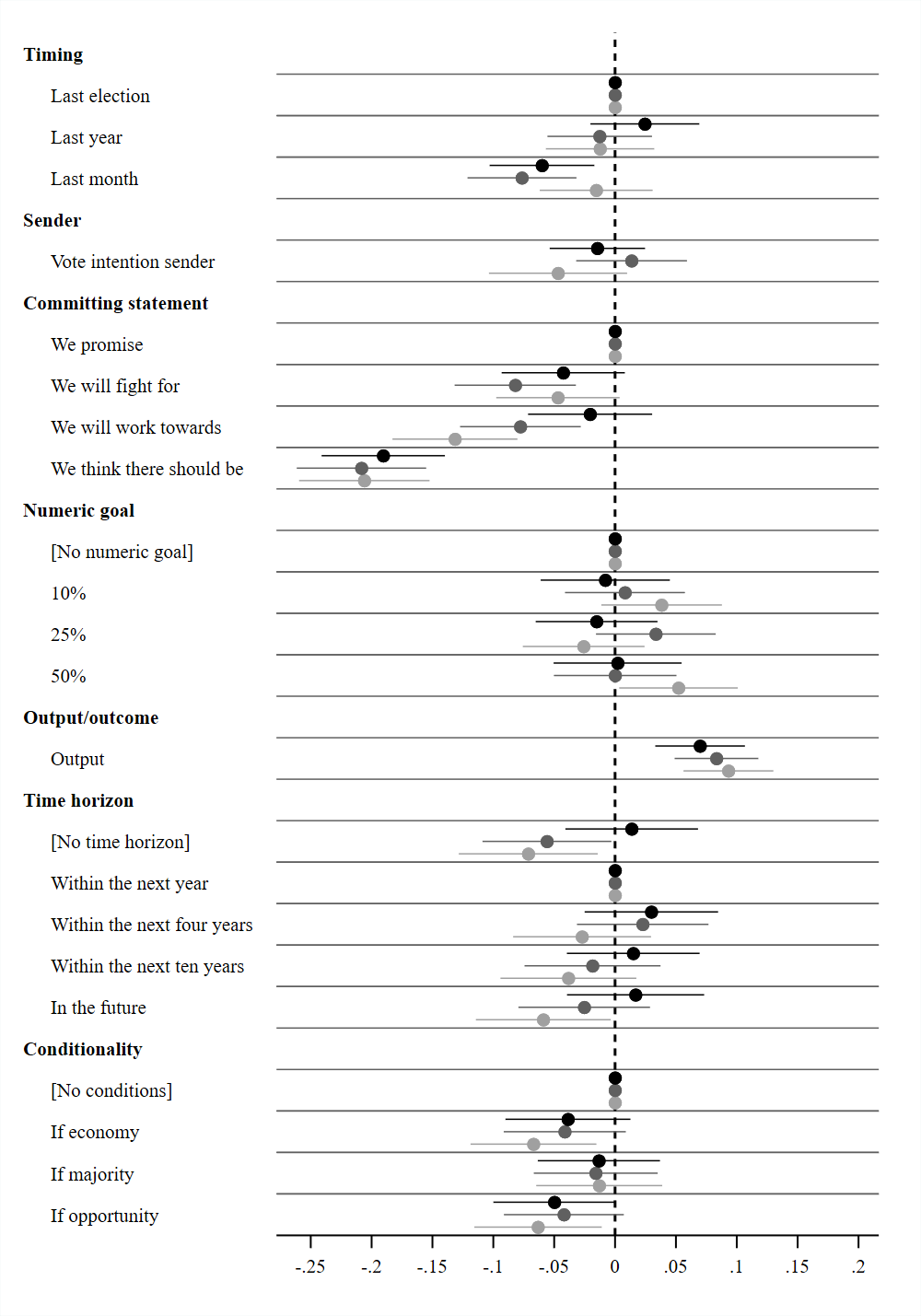
### Appendix D5d: Split-sample analysis with respondents below 39 years only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

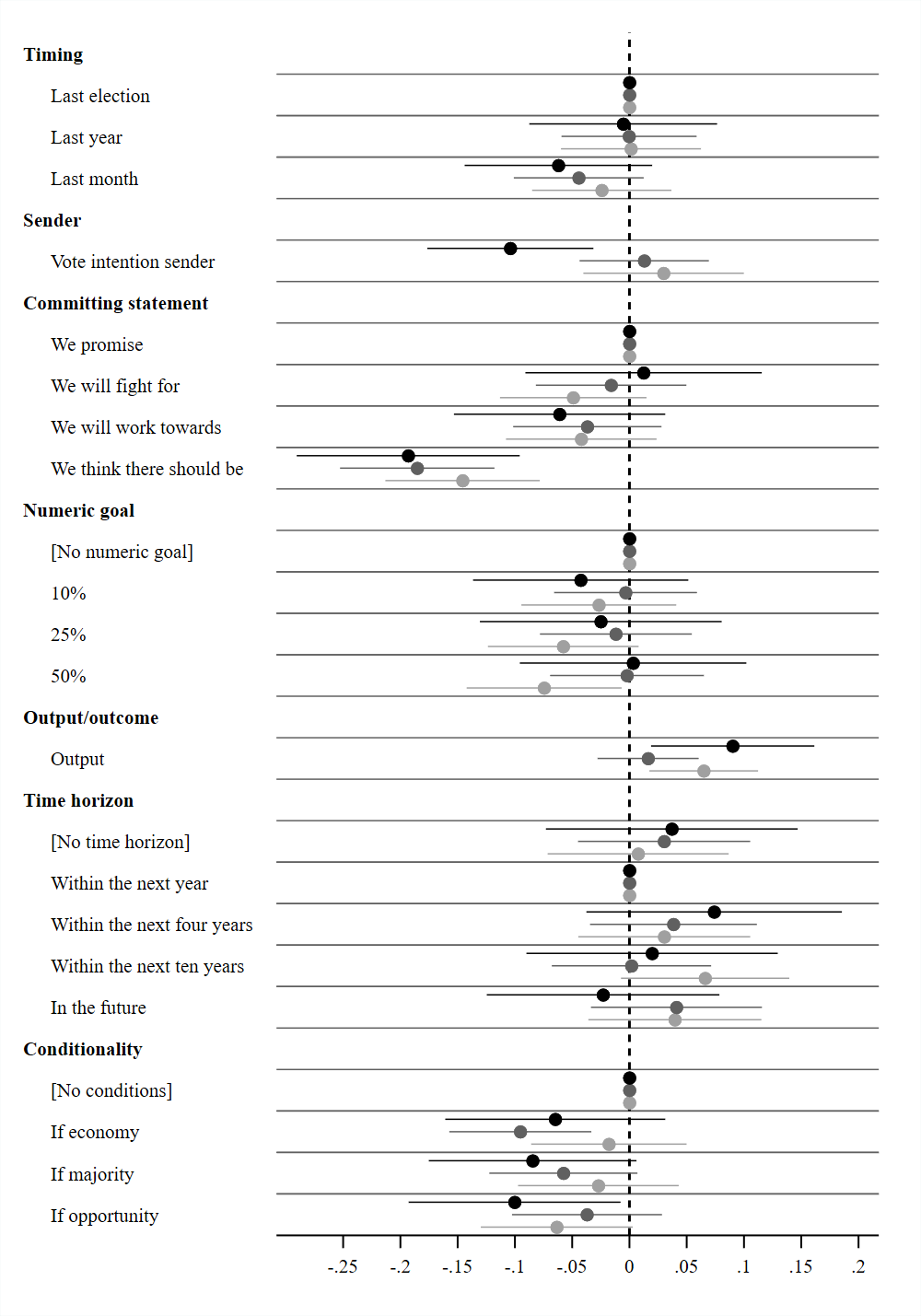
### Appendix D5e: Split-sample analysis with respondents between 40-65 years only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

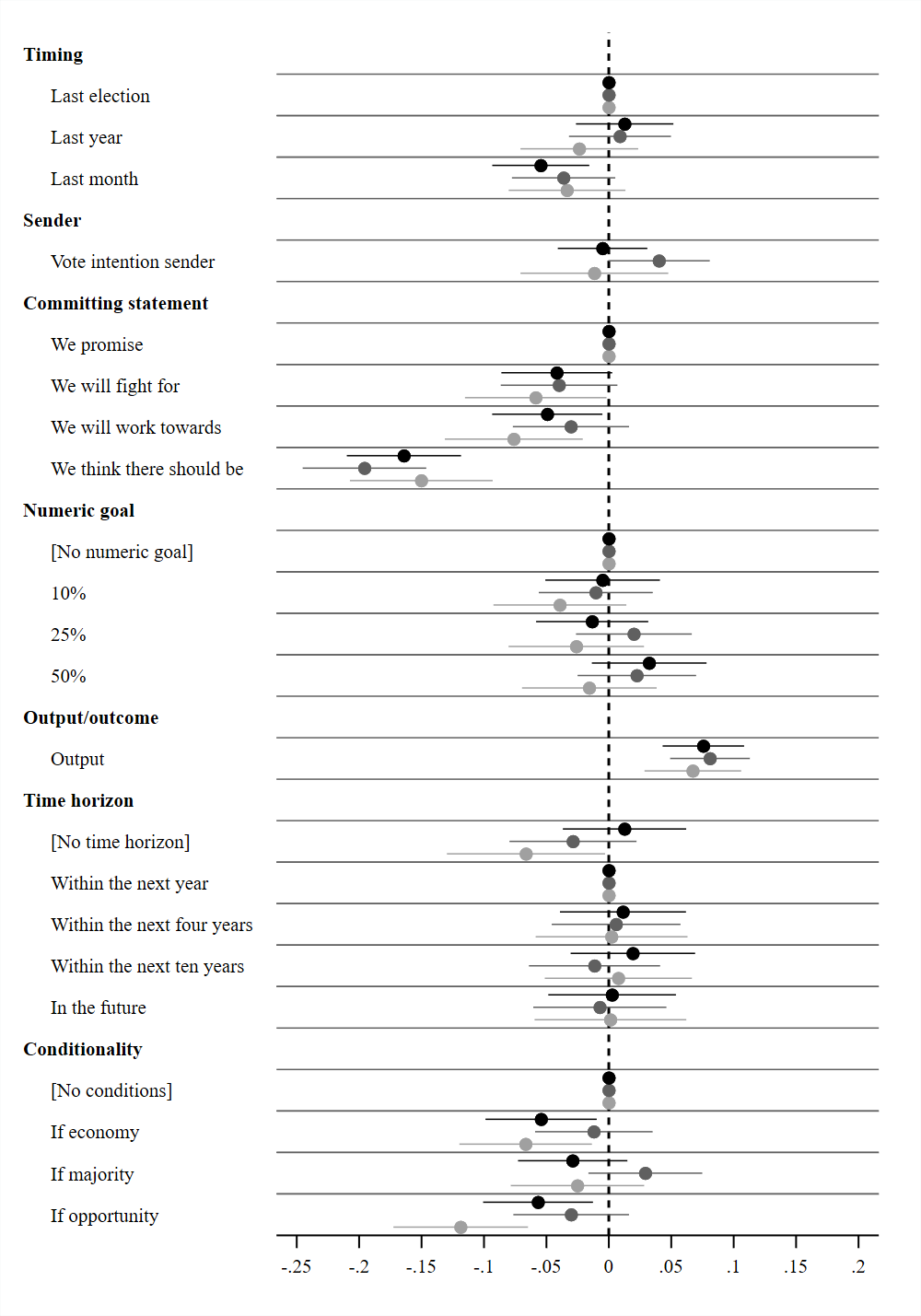
### Appendix D5f: Split-sample analysis with respondents above 65 years only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

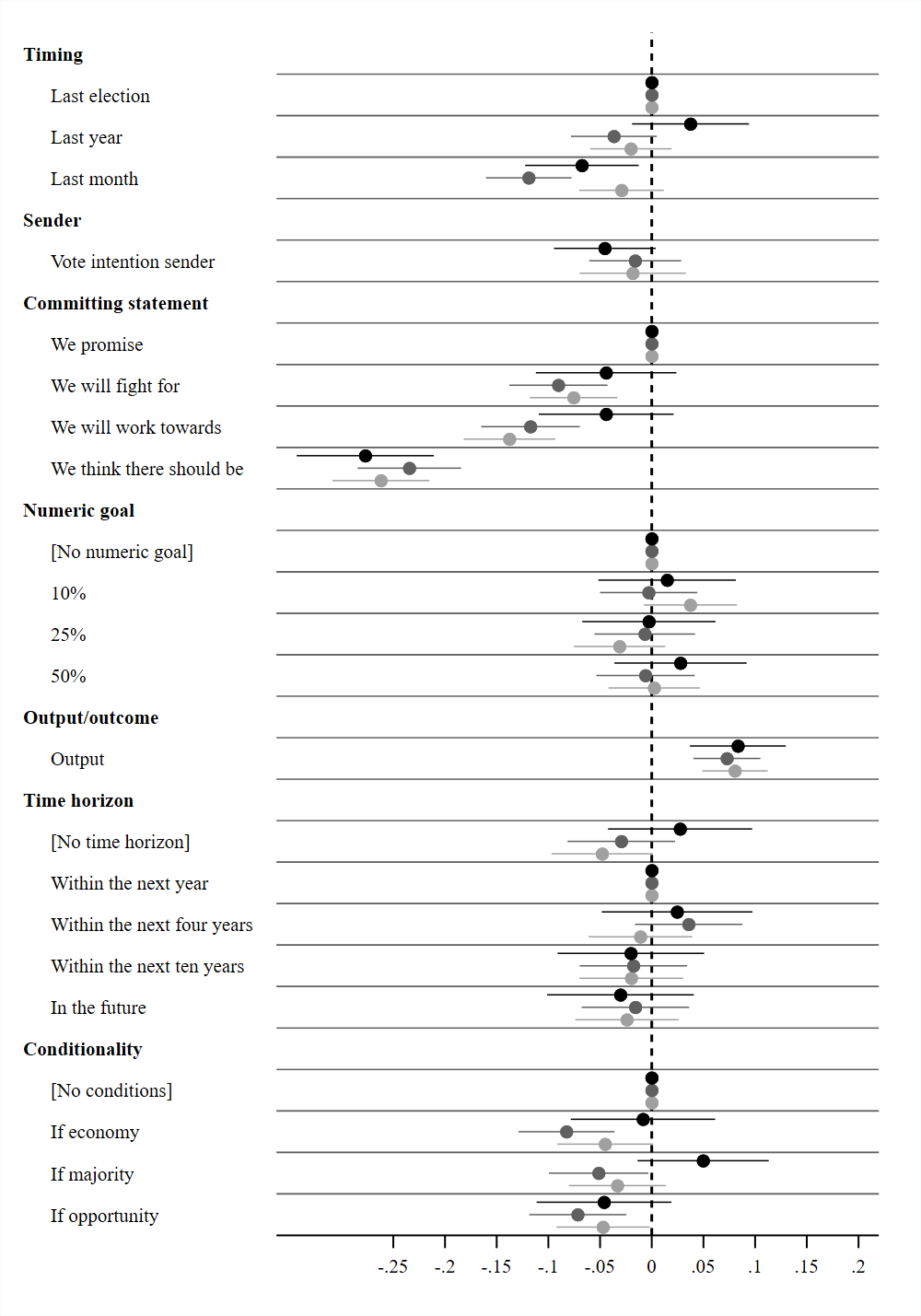
### Appendix D5g: Split-sample analysis with low educational level respondents only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

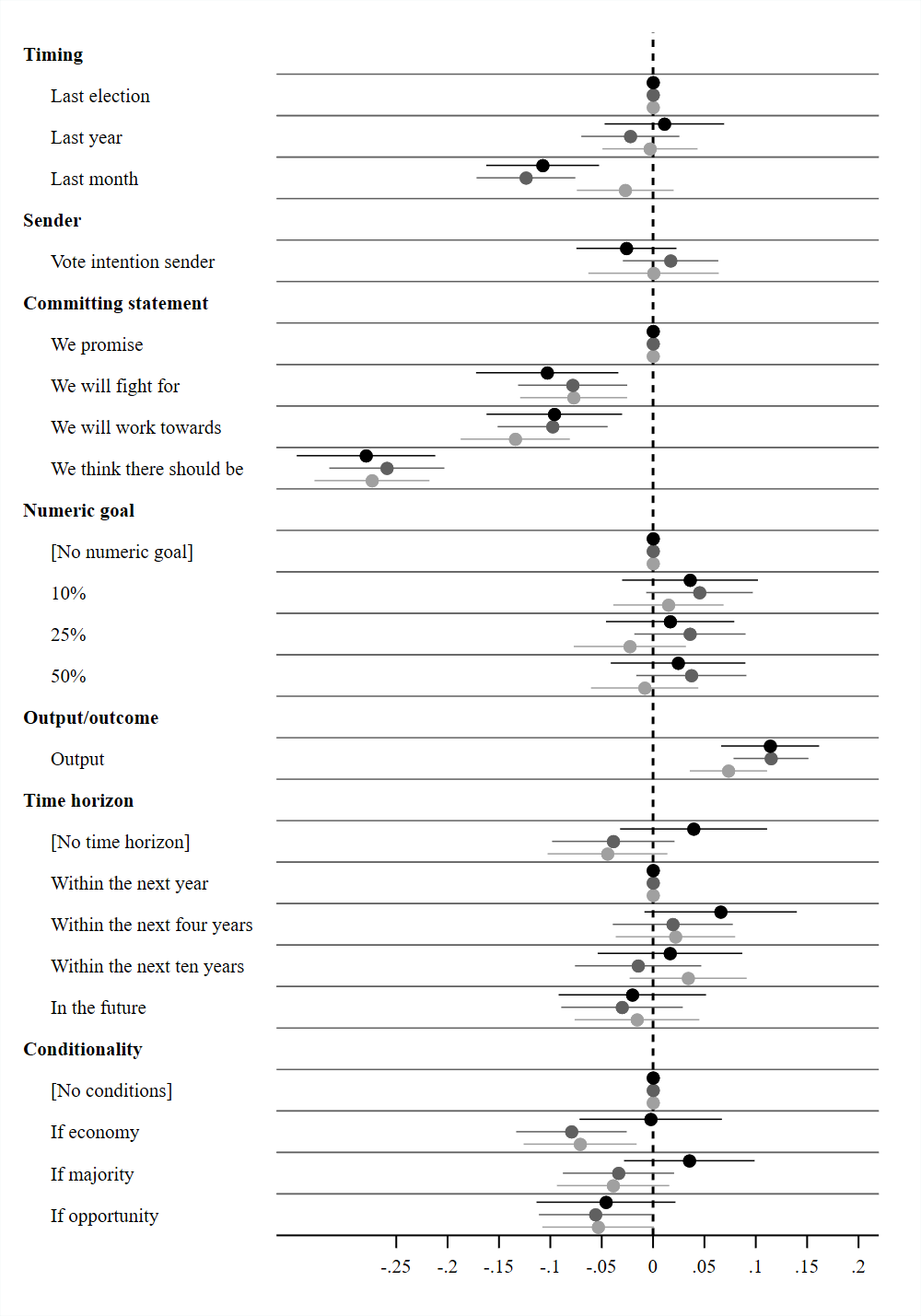
### Appendix D5h: Split-sample analysis with high educational level respondents only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

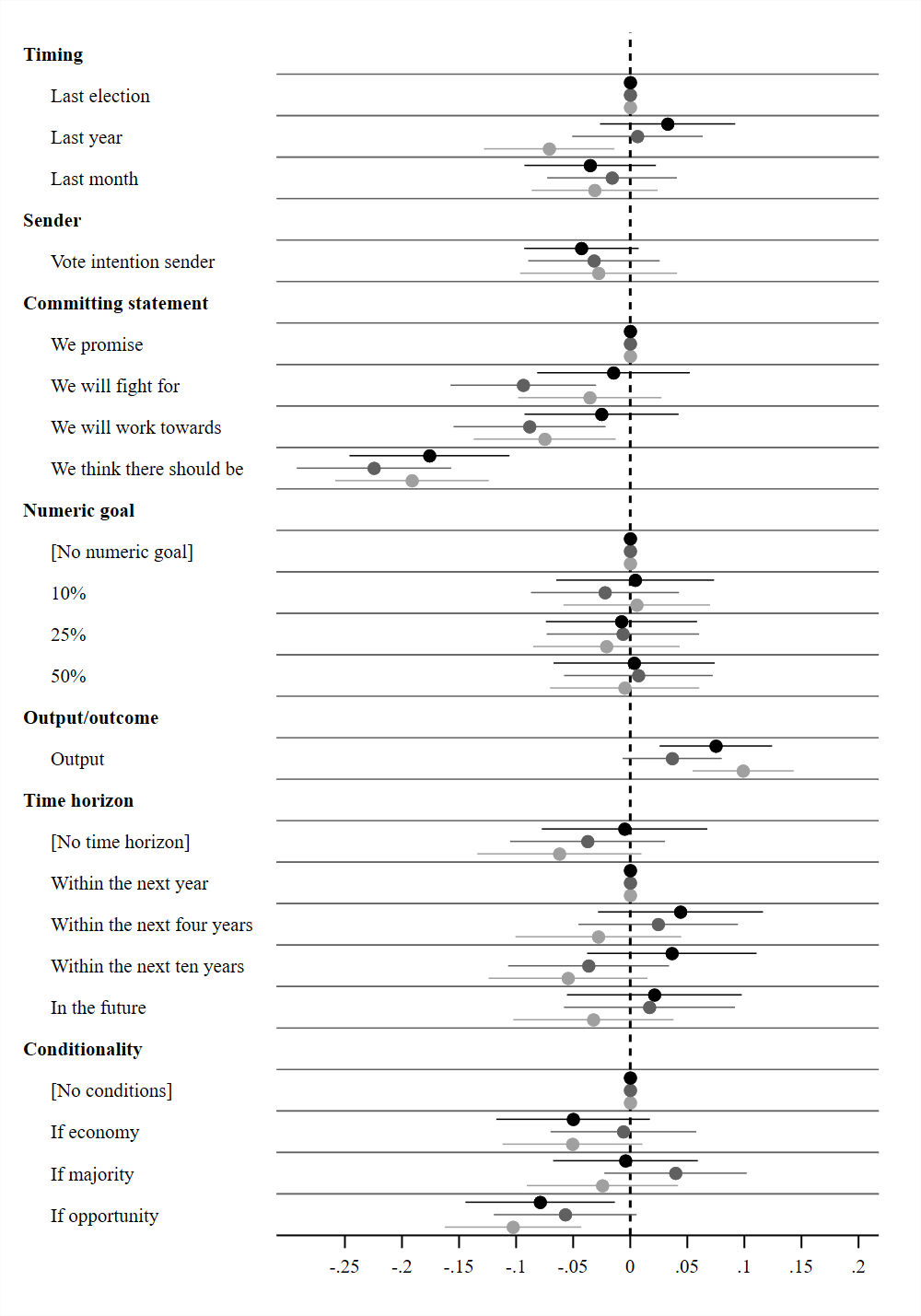
### Appendix D5i: Split-sample analysis with left-leaning respondents only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

### Appendix D5j: Split-sample analysis with right-leaning respondents only



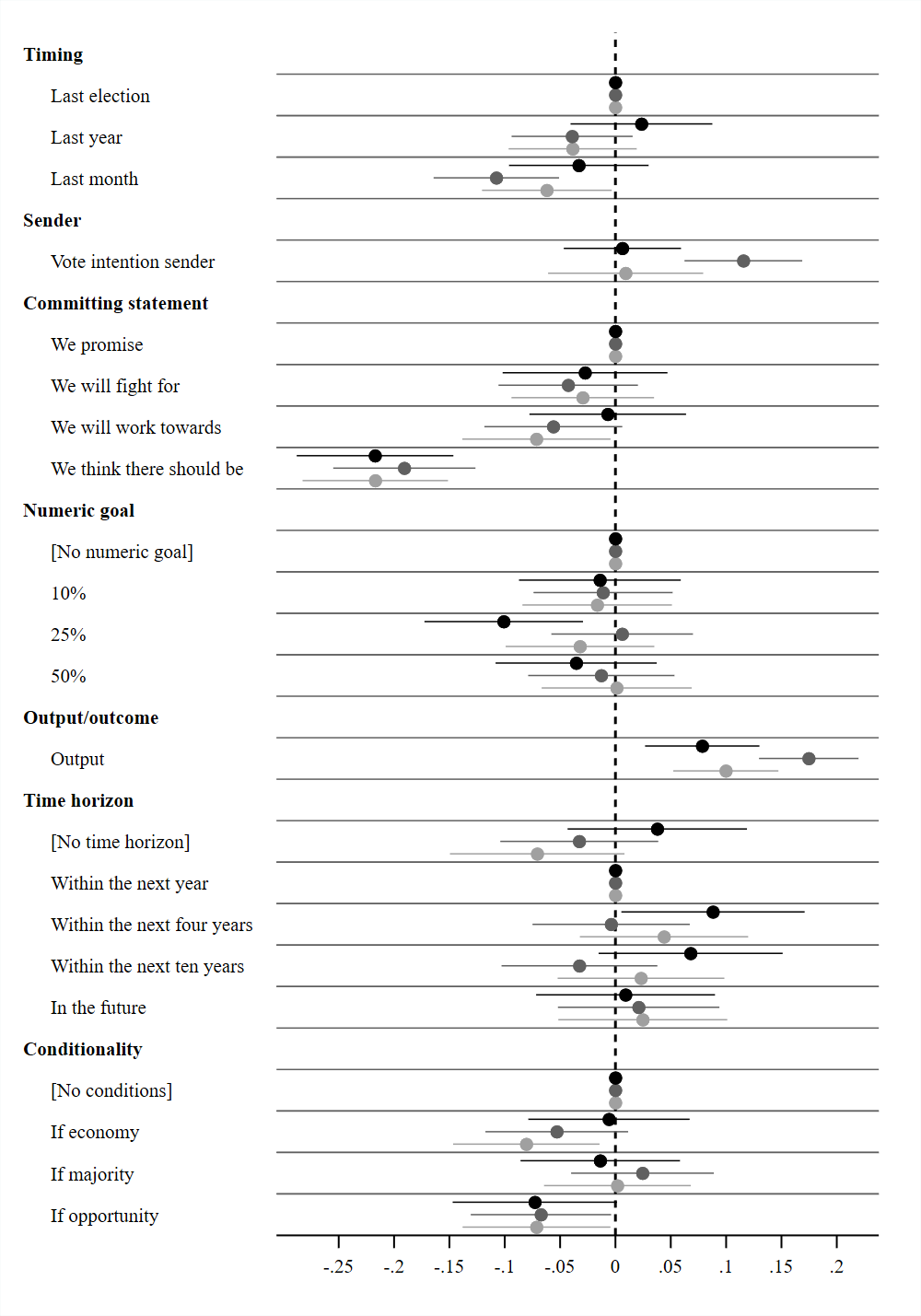
**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

## Appendix D6: Separate analyses across issues

This section examines whether the results are similar across issues. We do so by undertaking a range of split-sample estimations for each political issue: health, immigration, labor, and the Olympics. Again, although such estimations are generally quite inefficient, the results are quite similar across all four issues. The only clear exception is the output/outcome effect on immigration issues. Here, the effect is consistently insignificant in the United States and Denmark and even turns negative in Britain. We believe this is a consequence of our particular question formulations on the immigration issues—“fewer immigrants” (outcome) and “more resources to control family reunifications” (output). We suspect that the former formulation is too close to being an outcome: “resources” is a relatively vague description of an output and respondents may have focused on the purpose of the resources, namely to “control family reunifications,” which can be understood both as an output and an outcome. Whether this speculation is correct ultimately needs further analysis in order to draw conclusions with more certainty.

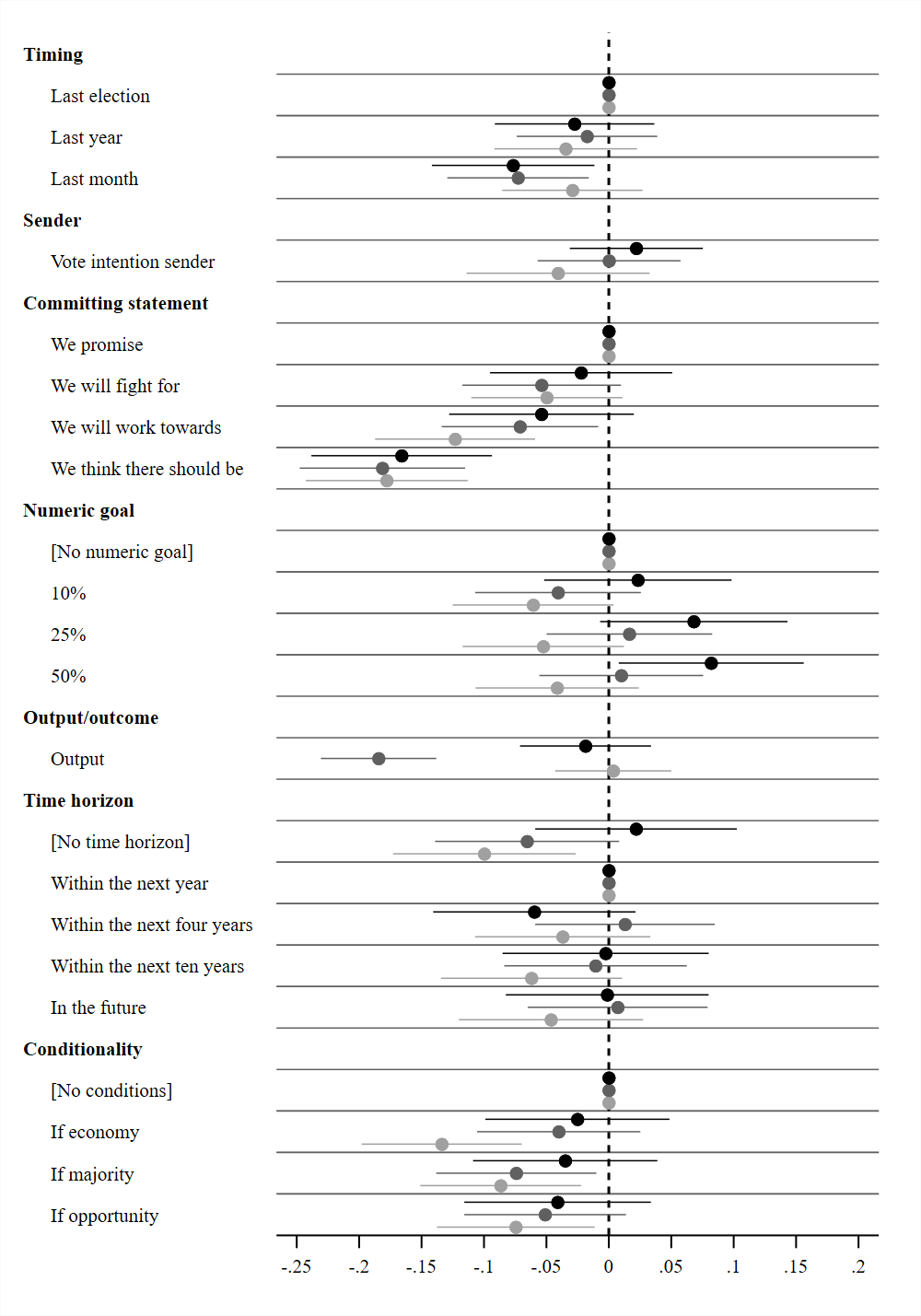
### Appendix D6a: Main models only for “Health” statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

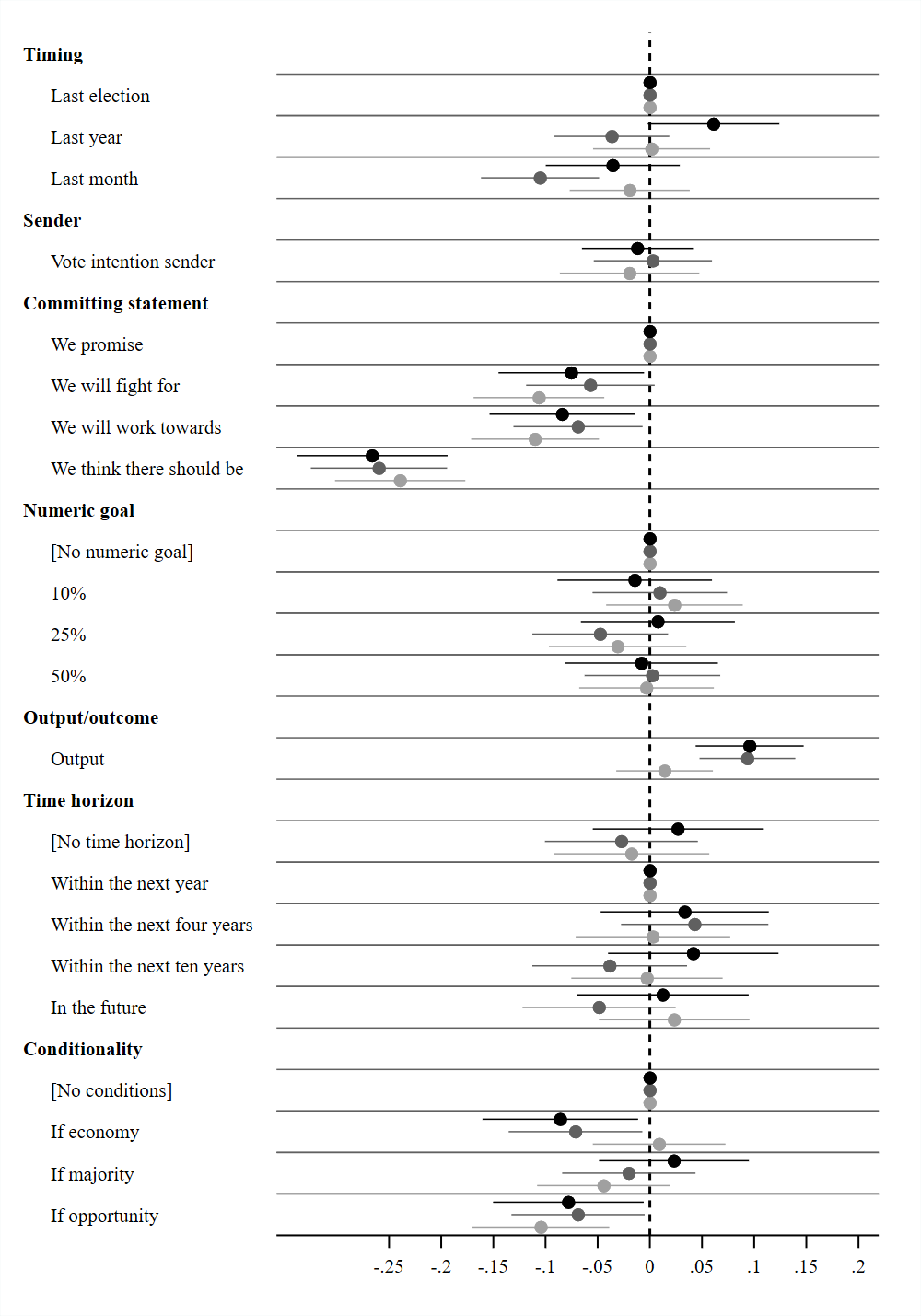
### Appendix D6b: Main models only for “Immigration” statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

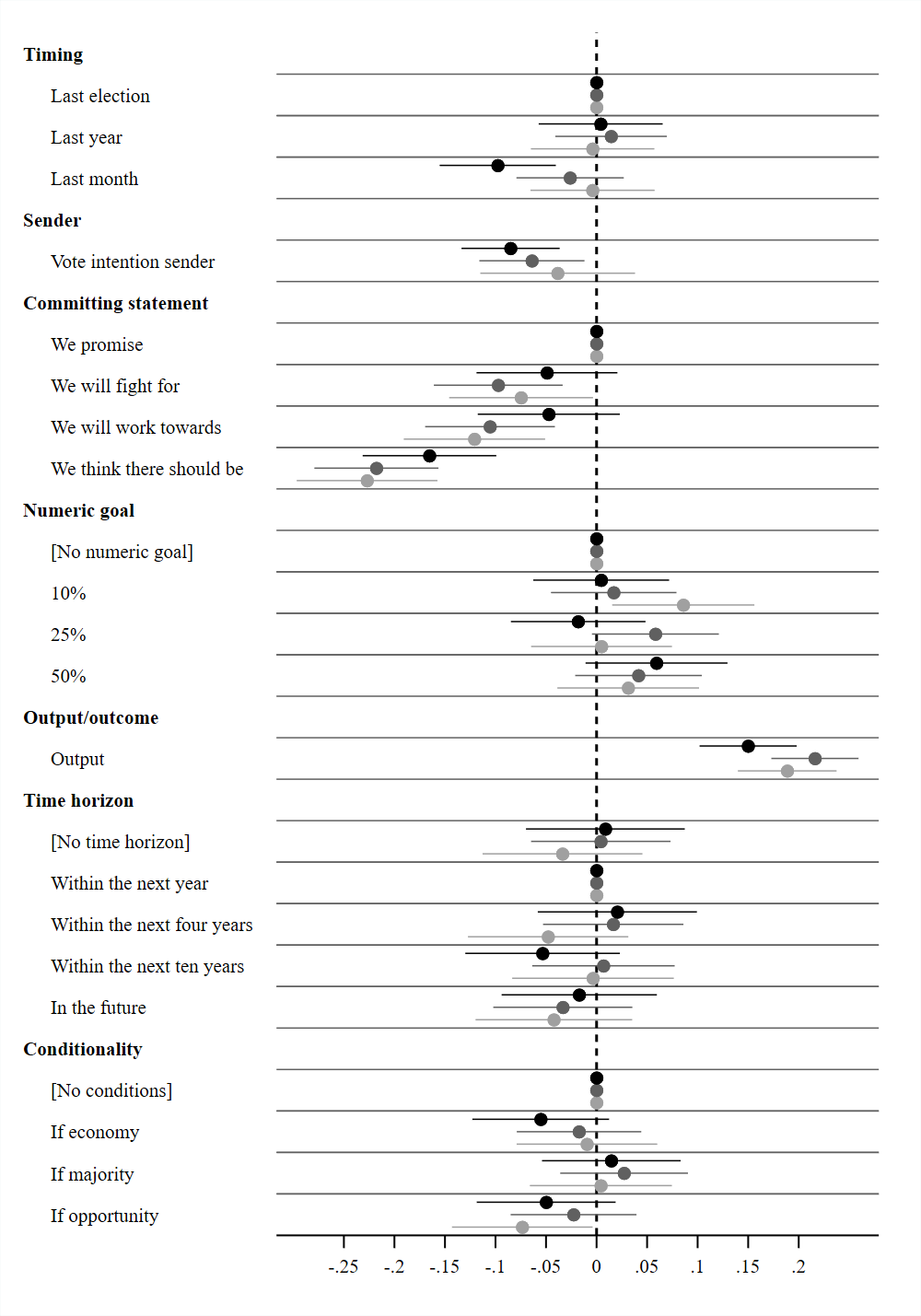
### Appendix D6c: Main models only for “Labor” statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

### Appendix D6d: Main models only for “Olympics” statements

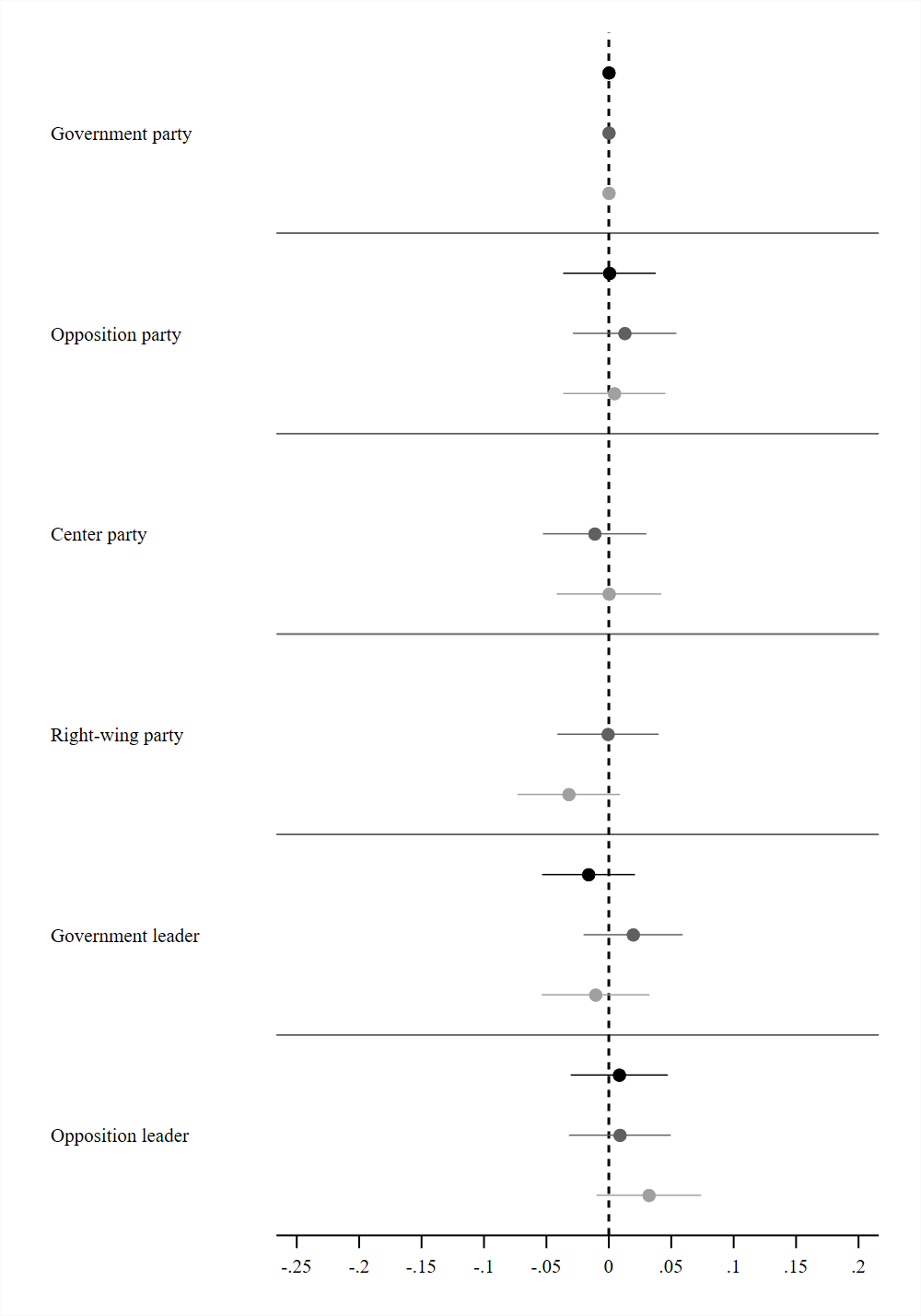


**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

## Appendix D7: Examining the direct effect of senders

In our main analysis, we scrutinize the effect of senders by examining whether a given respondent votes for the sender of a statement. In this section, we present the results of senders’ direct effect; that is, the effect of each sender regardless of whether the respondent votes for them or not. The results show that the sender has no direct effect and are thus in line with the results presented in the main article.



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

## Appendix D8: Examining ordering effects

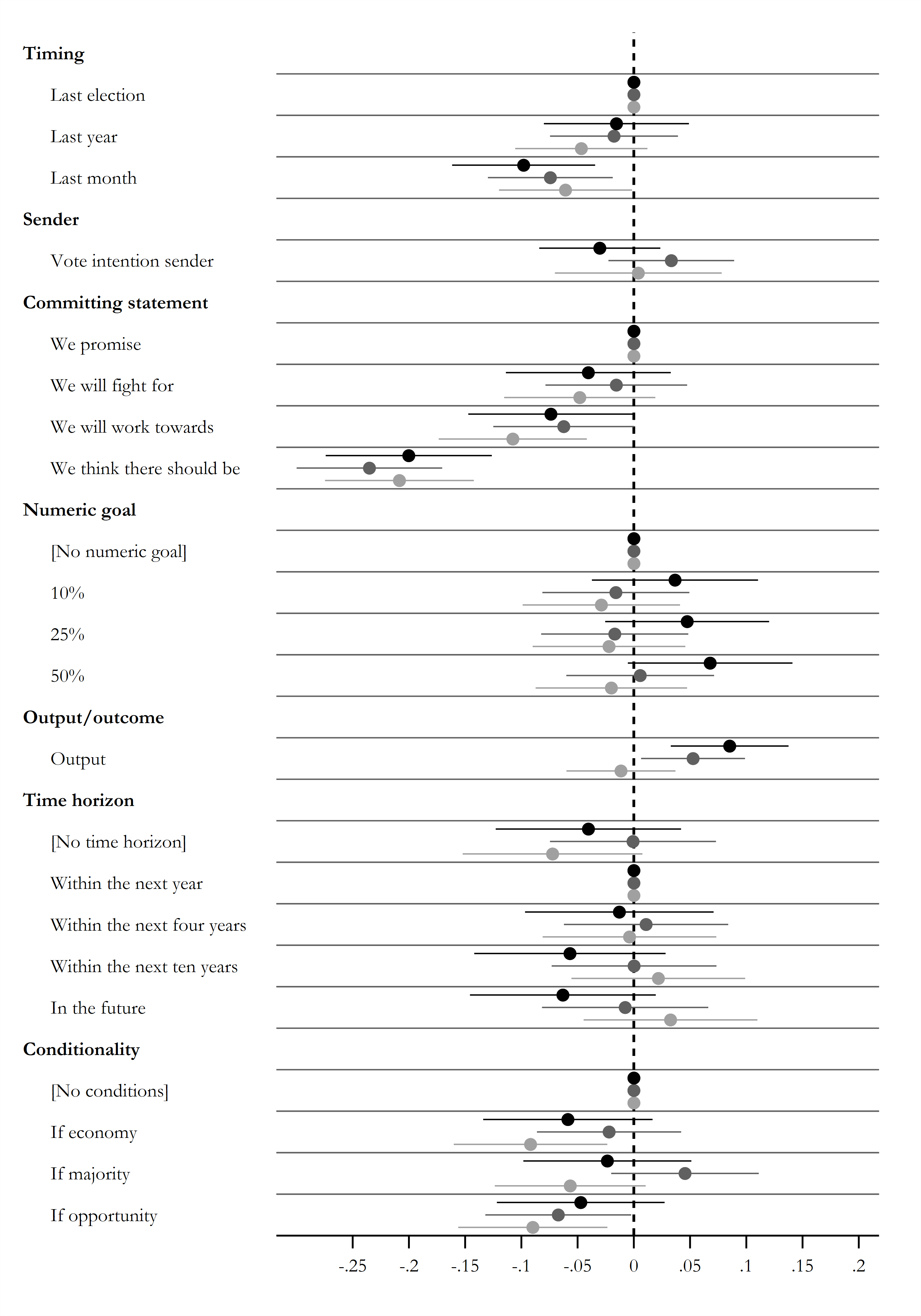
### Appendix D8a: Main results only for Round 1



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

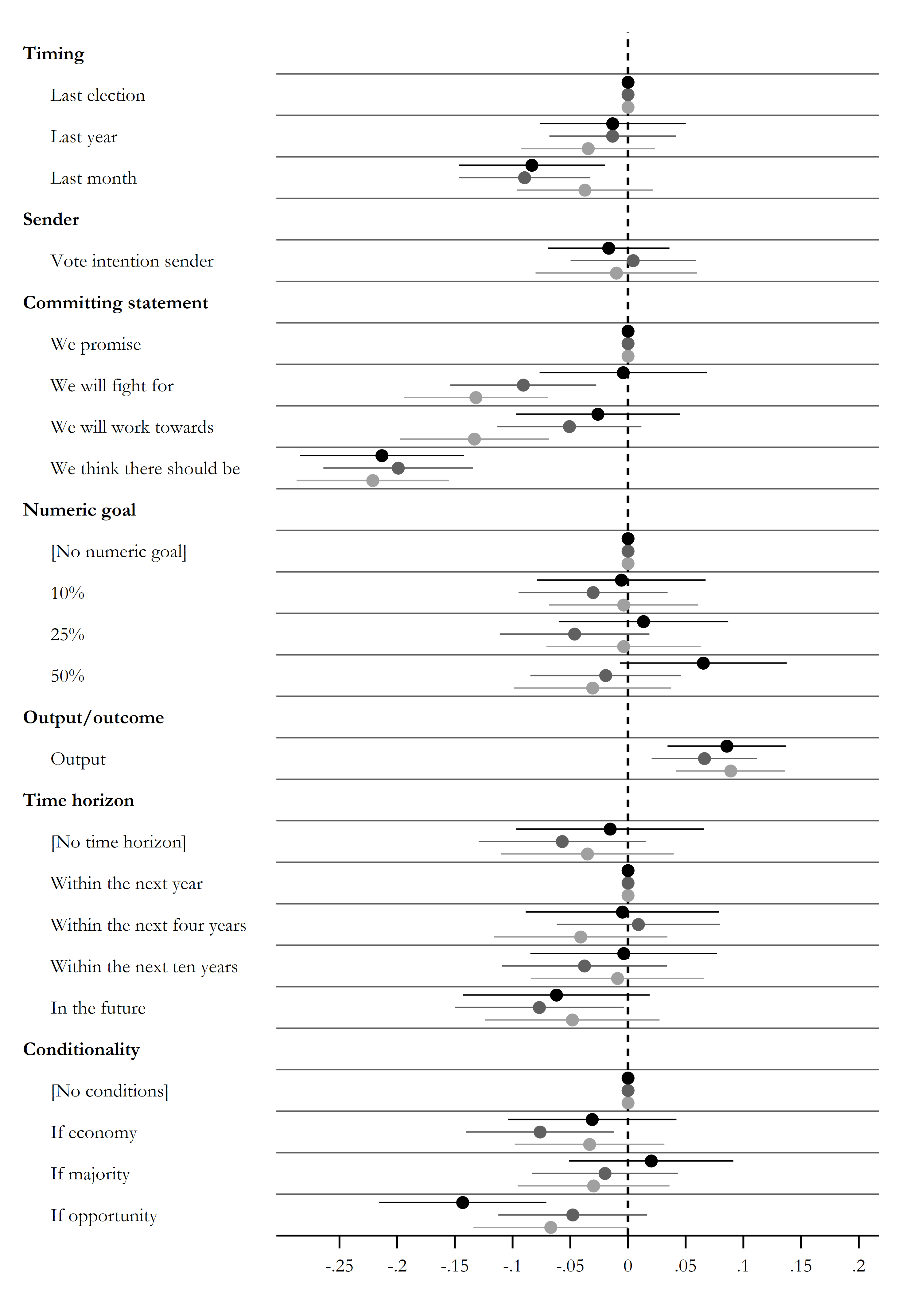
### Appendix D8b: Main results only for Round 2



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

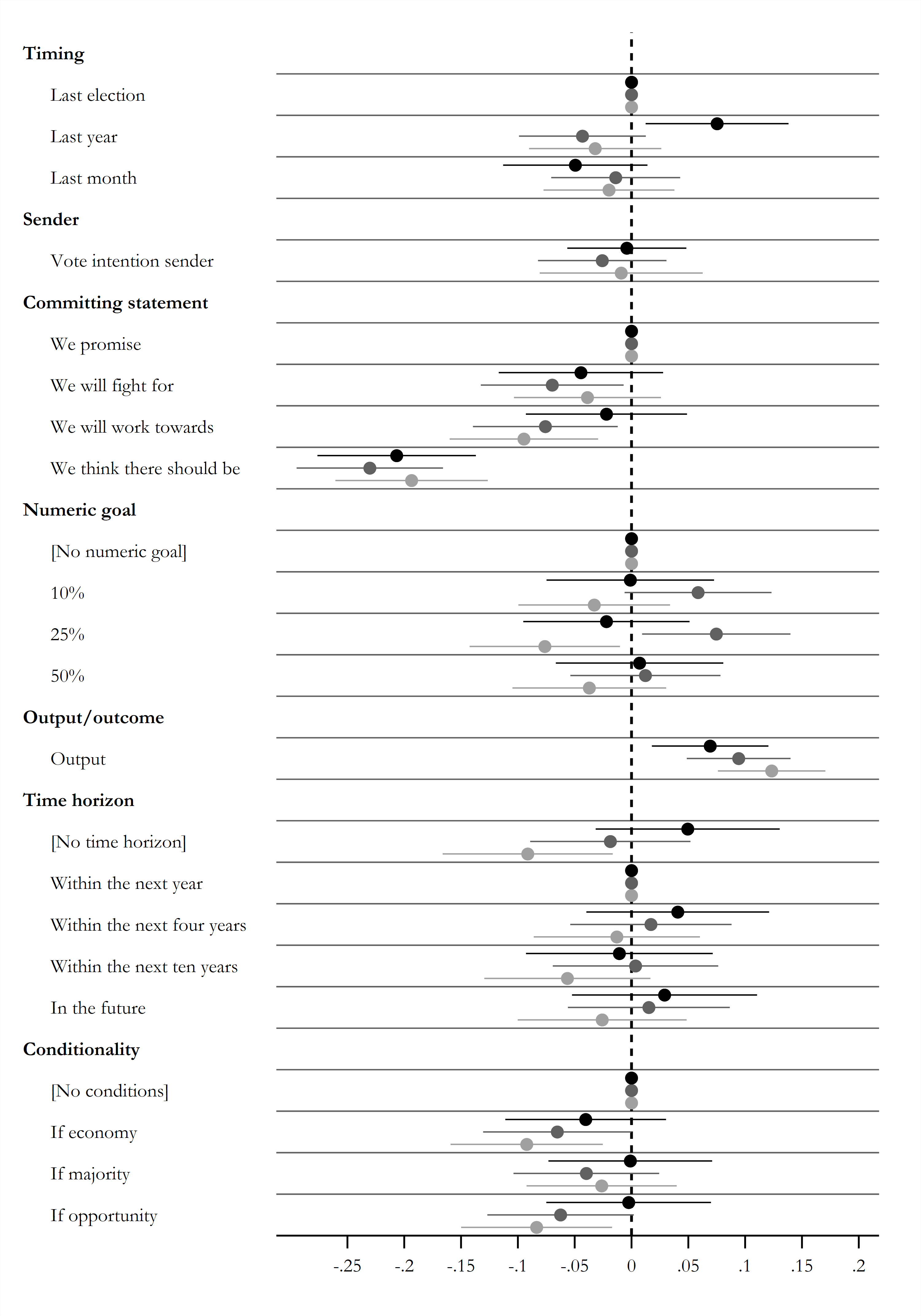
### Appendix D8c: Main results only for Round 3



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

### Appendix D8d: Main results only for Round 4



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

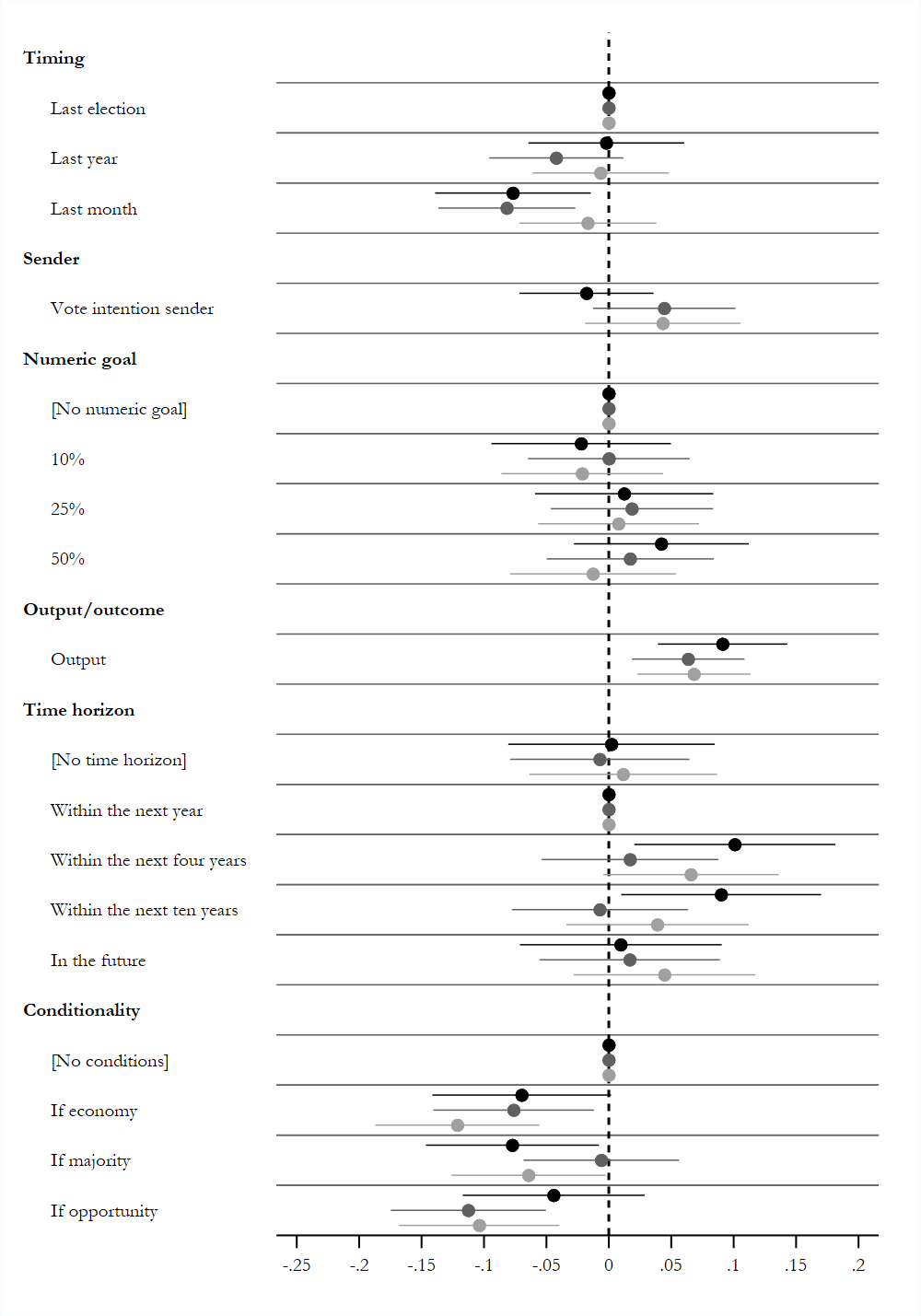
Black = United States, dark grey = Britain, and light grey = Denmark.

## Appendix D9: Examining interaction between attributes through split-samples

In this section, we further assess the assumption that the effects of each attribute are independent of other attributes. To do this, we run split-sample analyses conditioning the models on each significant attribute: statements with “we promise” only, “we think there should be” only, no conditions only, “if opportunity” conditions only, and outputs only.

The confidence intervals are naturally larger due to the restricted sample sizes, but overall the conclusions are similar: The attributes that show significant versus insignificant effects are generally the same. The only exception is that the conditionality attributes exert insignificant effects when the committing statements are held constant at “we think there should be.” Yet this is understandable given that we expect these two attributes to be part of the same perceptual dimension, namely *commitment*. Using a very weak committing formulation already weakens the degree of commitment, which naturally weakens the impact of attaching further conditions. Apart from this, all other coefficients yield similar conclusions.

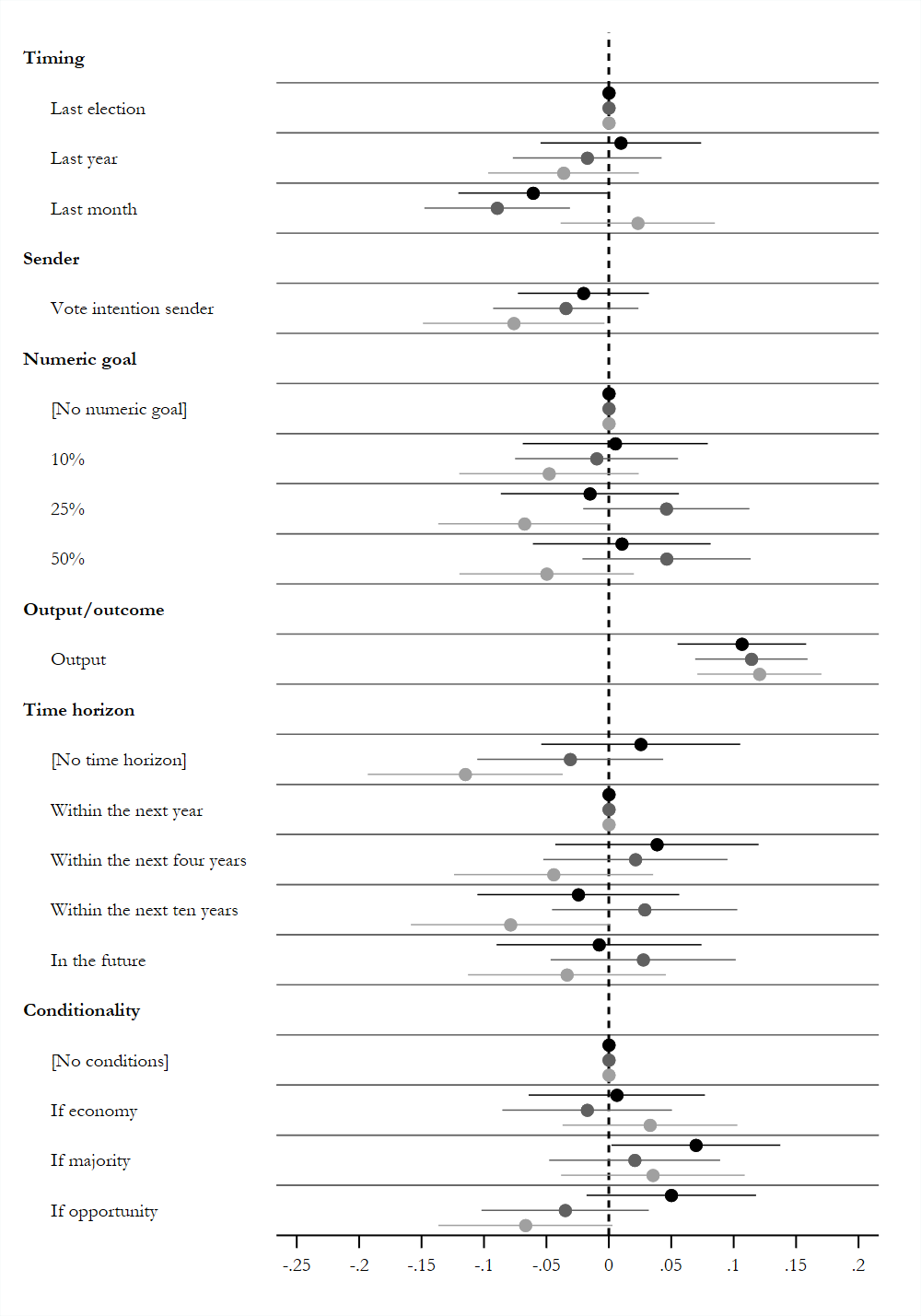
### Appendix D9a: Main models only for “We promise” statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

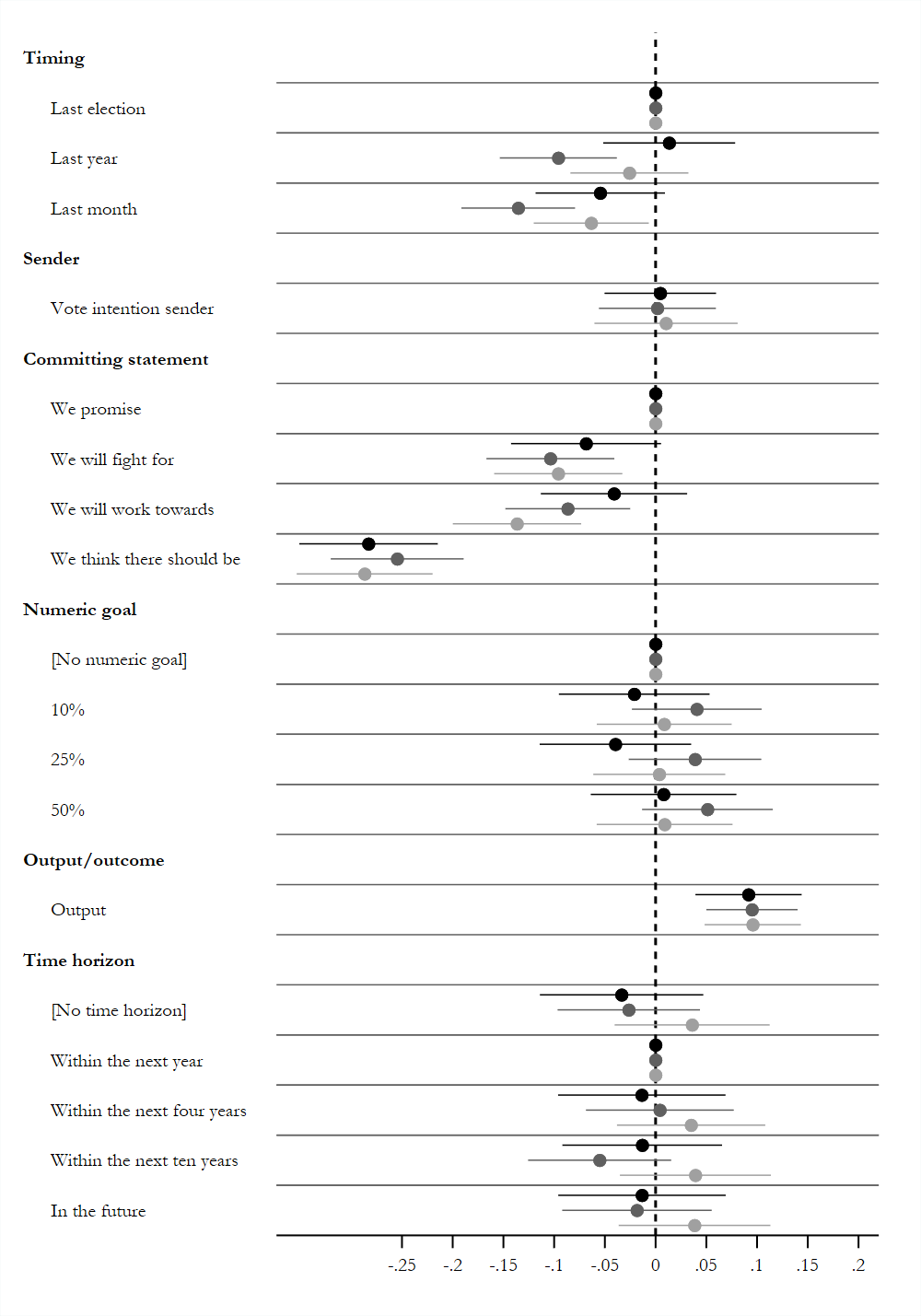
### Appendix D9b: Main models only for “We think there should be” statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

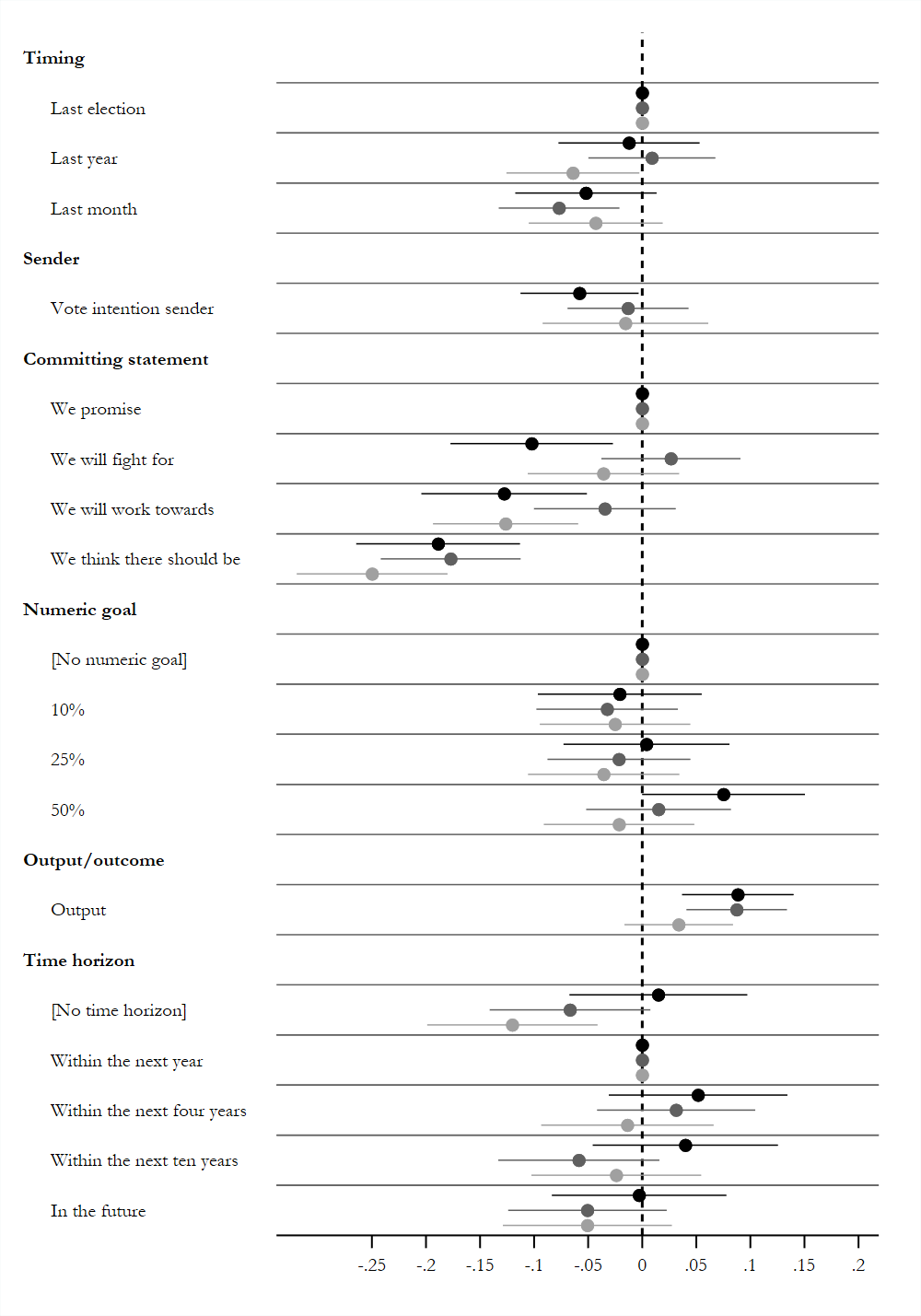
### Appendix D9c: Main models only for no conditions statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

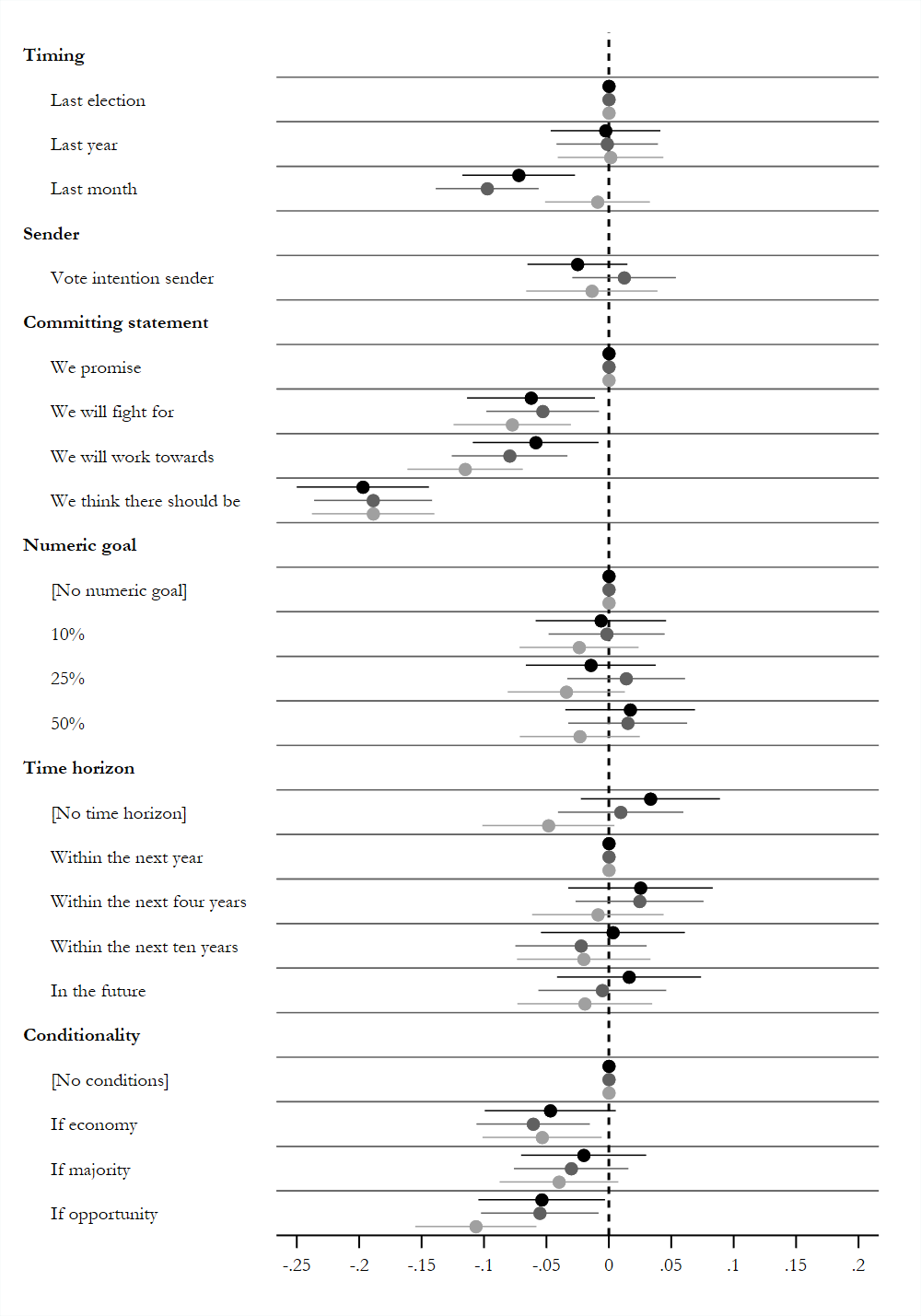
### Appendix D9d: Main models only for “if opportunity” statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

### Appendix D9e: Main models only for output statements



**Note:** Average marginal component effects (AMCE) on the likelihood of considering a statement an election pledge and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

# Appendix E: Mediation analysis

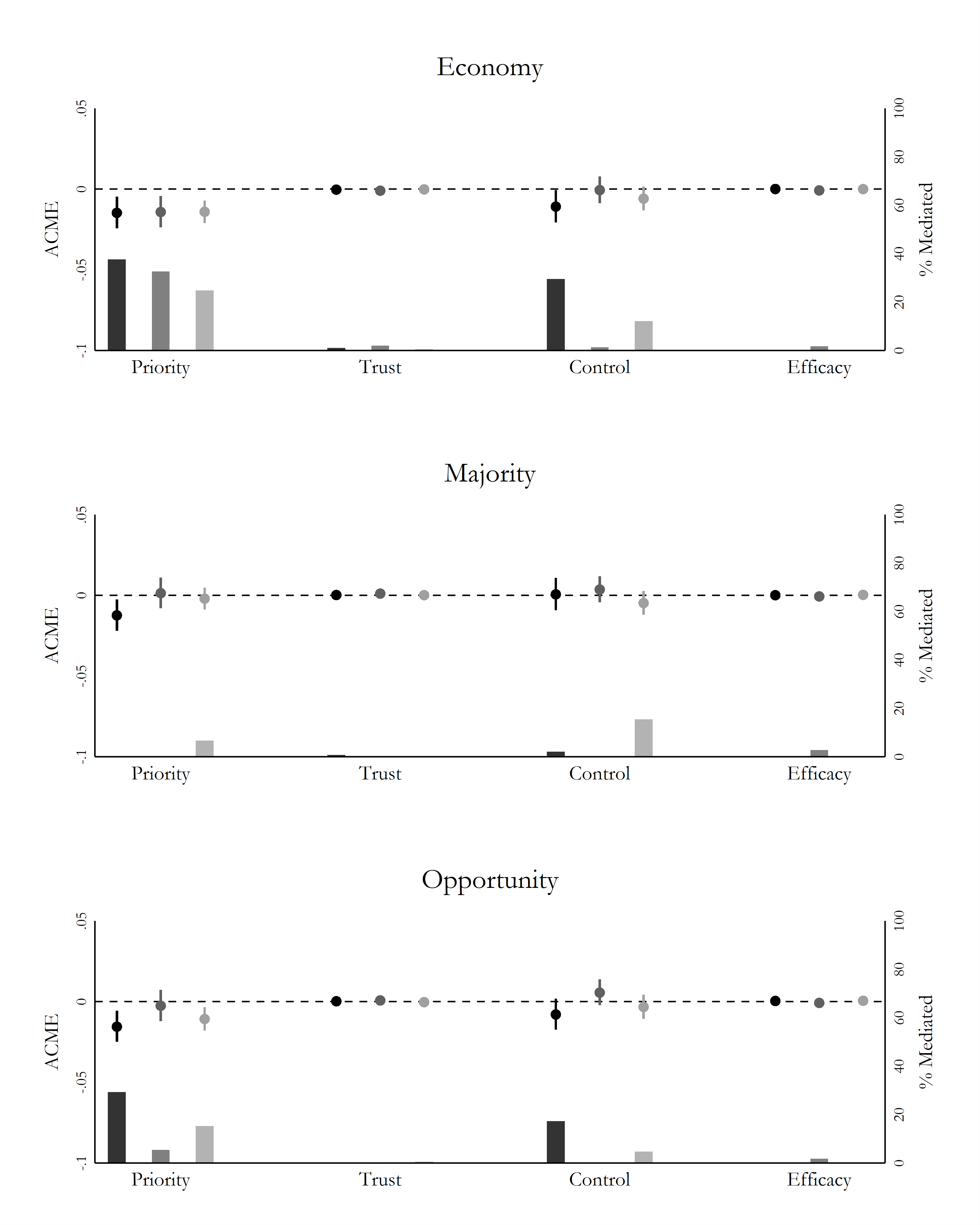
In this section, we present the mediation analyses in more detail. We use the causal mediation analysis approach stipulated by Imai and Yamamoto (2013) in order to examine the mechanisms between different statements and pledge categorization. Specifically, we utilize the fact that we inquired about respondents’ perceptions of sender priority,[[1]](#footnote-1) sender trustworthiness,[[2]](#footnote-2) sender control,[[3]](#footnote-3) and respondent efficacy[[4]](#footnote-4) after each statement. That is, we measured respondents’ perceptions of each of the four mechanisms summarized in Table 1 in the main article. The causal mediation approach then enables a test of whether the effects of the different statement attributes, analyzed in the main analysis, affect respondents’ pledge categorization through any of these four mechanisms. We acknowledge that this approach is based on assumptions that are often violated in many settings, but it nonetheless provides correlational illustrations of our argument.

Specifically, we present the results of all significant attributes; that is, *committing statement* (the mediation effects of “we think there should be” compared to “we promise” as reference category), *conditionality* (the mediation effects of “if the economy allows it,” “if a majority in Congress/Parliament is in favor,” and “if the opportunity arises” compared to not attaching a condition as reference category), as well as *output/outcome* (the mediation effects of outputs compared to outcomes as reference category).

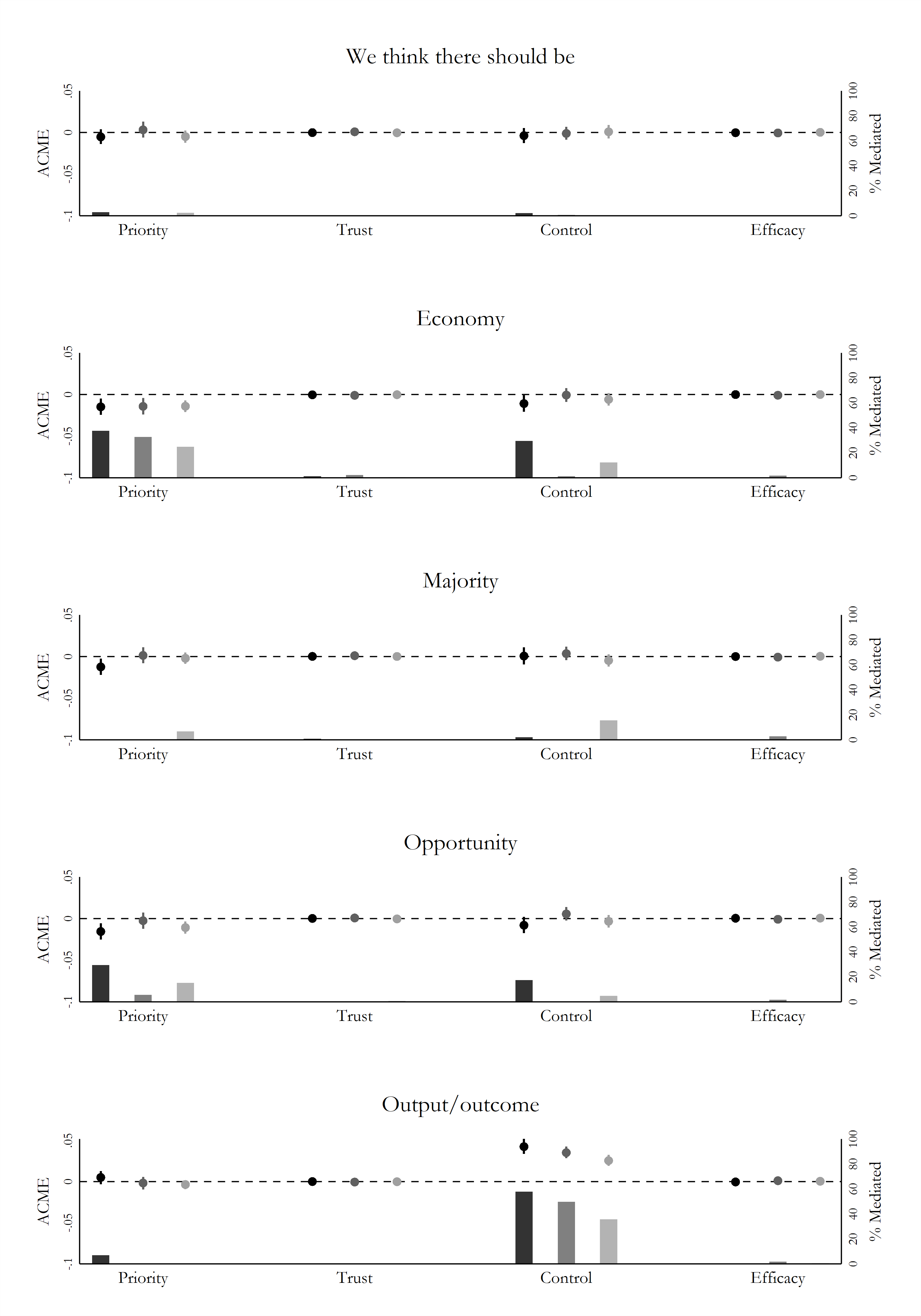
In general, we recover the mediation effects as expected. Still, some mediation effects are not as expected. First, the mediation effect of attaching the condition “if a majority in Congress/Parliament is in favor” shows similar patterns, but fails to attain significance in Britain and Denmark. Yet, given that the main effects of this particular attribute were also not robust (see main article), it is no surprise that the mediation effects are weak as well.

Secondly, we do not find the expected mediation effects of the committing statement. In short, even though a committing statement is shown to be a strong and significant determinant of whether voters consider a statement a pledge, it does not seem to matter because of voters’ perceptions of how much the politician prioritizes a given issue. Perhaps this is due to the fact that a committing formulation is so fundamental in terms of perceiving (or not perceiving) a statement to be a pledge that it matters for many reasons—not just in terms of priority. A pledge that does not contain an explicit commitment, such as “we promise,” “we vow,” or “we swear,” is hard to meaningfully label a pledge at all for many reasons, and therefore, while the committing statement exerts strong effects overall on pledge perceptions, it does not seem to be mediated by one particular factor.

## Appendix E1: Mediation analysis of conditionality attributes



## Appendix E2: Mediation analysis of all significant attributes

****

# Appendix F: Accountability analysis

In this section, we present the accountability analyses in more detail. First, we present cross-tabulations of perceptions of election pledges and perceptions of accountability. Next, we provide some robustness checks that in different ways seek to account for whether the sender is a government party or not, and whether the respondent has voted for a government party or not.

## Appendix F1: Cross-correlation of election pledge and accountability

Do voters think that politicians who make election pledges should be held accountable for those pledges? In general, this seems to be the case. If a statement is categorized as a pledge, an average of 80% of voters believe the pledge giver is accountable for implementing it. If a statement is not seen as a pledge, an average of 66% do not hold the pledge giver accountable. Thus, voters seem to perceive pledges as a distinct form of political communication—both in terms of their rhetorical quality and their associated accountability.

Voters’ Perceptions of Election Pledges and Accountability

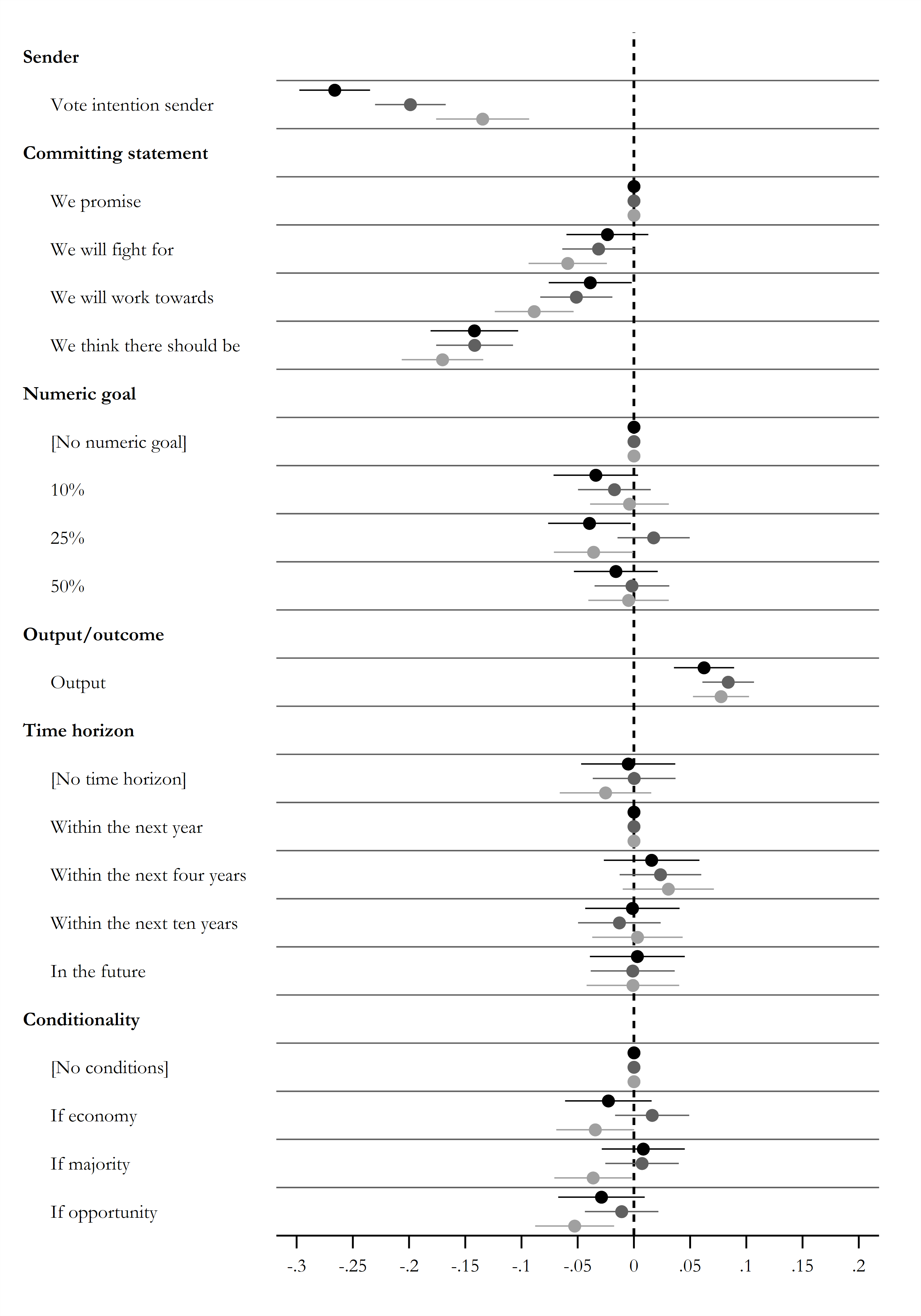
|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Should [sender] be held accountable?** | |
| **Is the statement an election pledge?** |  | **Yes** | **No** |
| **Yes** | *United States* | 77% | 23% |
| *Britain* | 83% | 17% |
| *Denmark* | 80% | 20% |
| **No** | *United States* | 31% | 69% |
| *Britain* | 41% | 59% |
| *Denmark* | 30% | 70% |

**Notes:** Entries denote the percentage of responses across rows for each country.

## Appendix F2: Controlling for party choice

In this section, we include a categorical variable controlling for sending party and whether the sending party is in government, respectively. This is done in order to assess whether the significant results of voting for sender on accountability perceptions is confounded by the party identification itself or the sender being in government or not. The results of all these additional tests yield the same conclusion as in the main article.

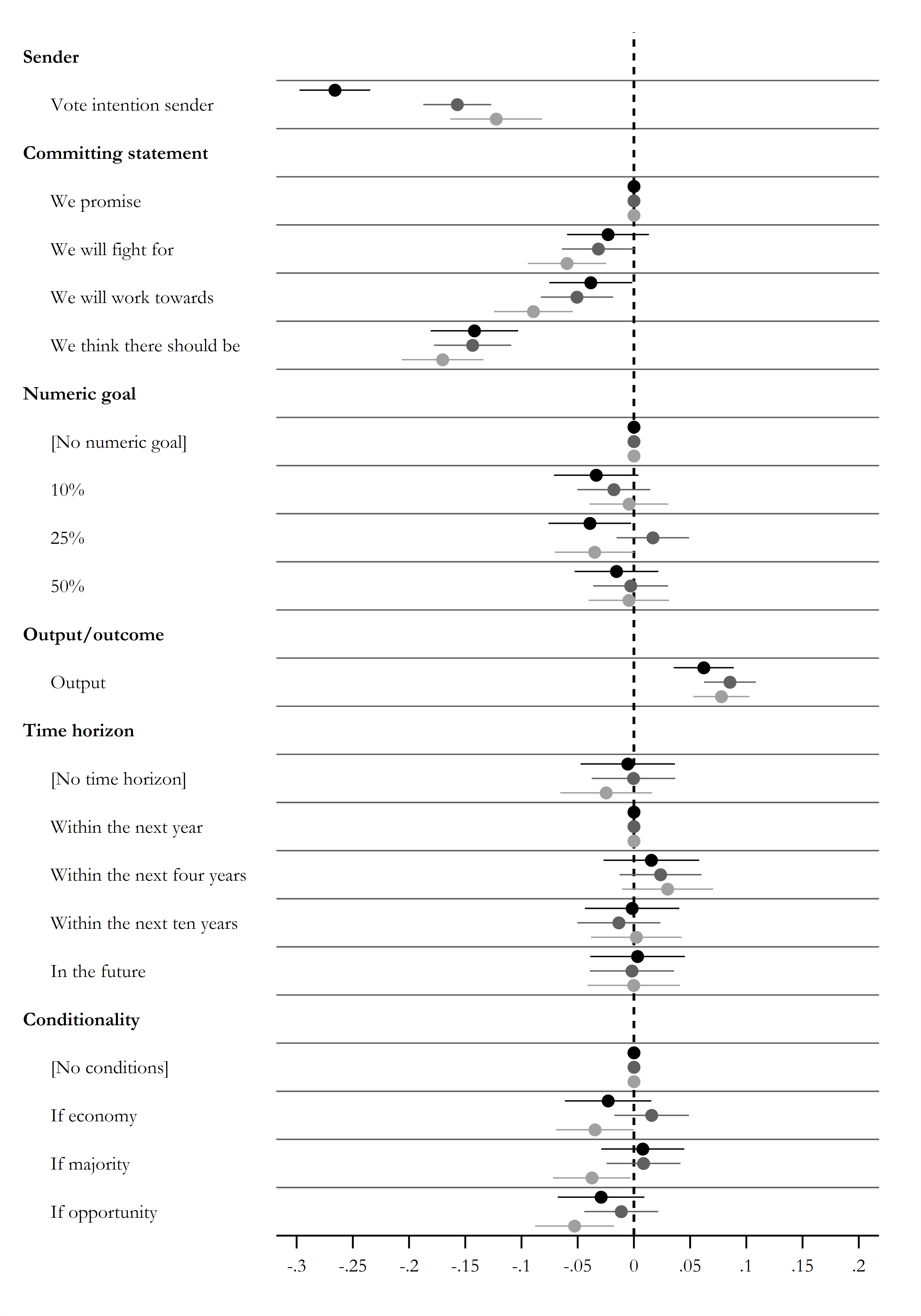
### Appendix F2a: Controlling for sender



**Note:** Average marginal component effects (AMCE) on the likelihood of considering that sender should be held accountable and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

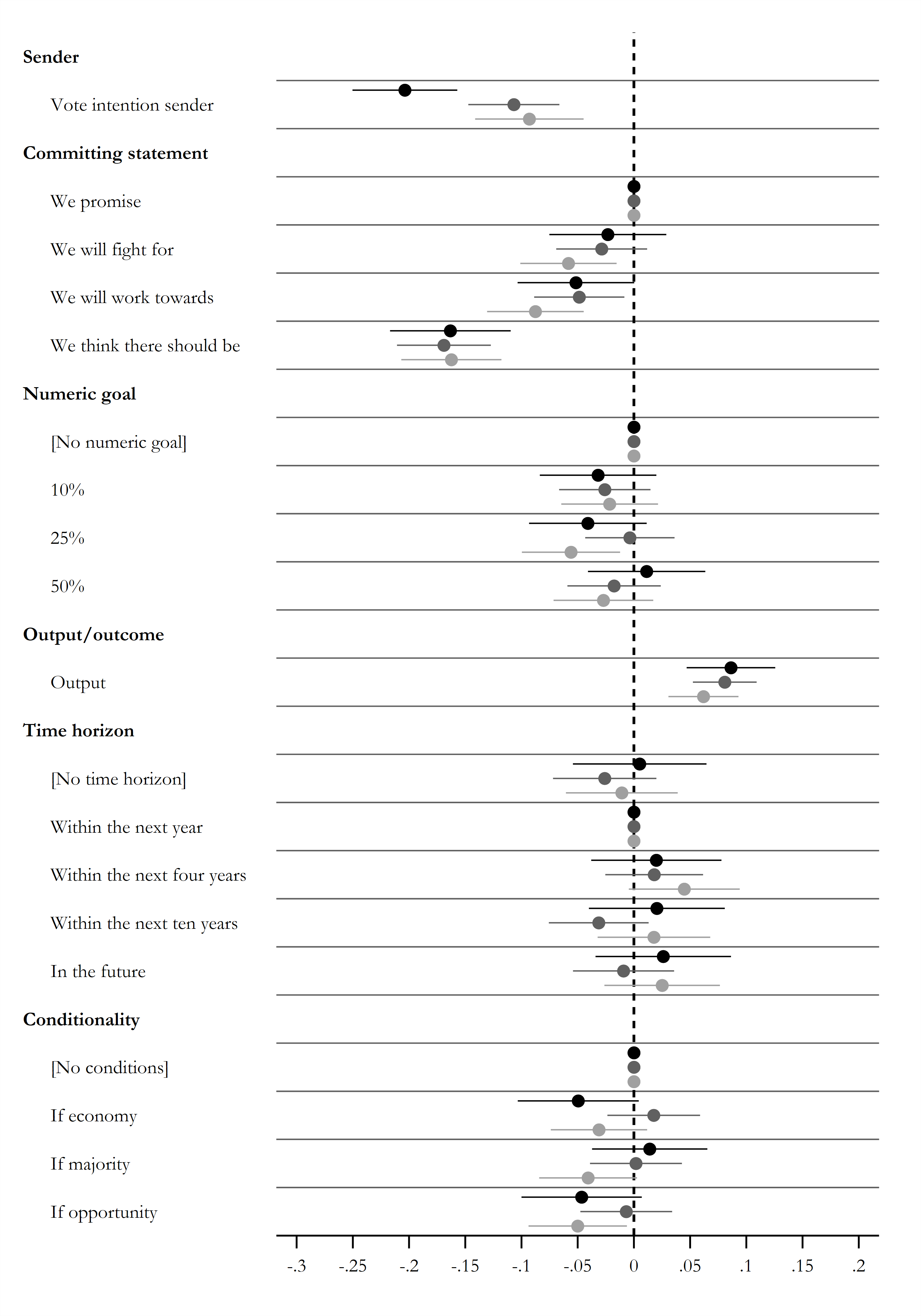
### Appendix F2b: Controlling for sender’s status as in government or opposition



**Note:** Average marginal component effects (AMCE) on the likelihood of considering that sender should be held accountable and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

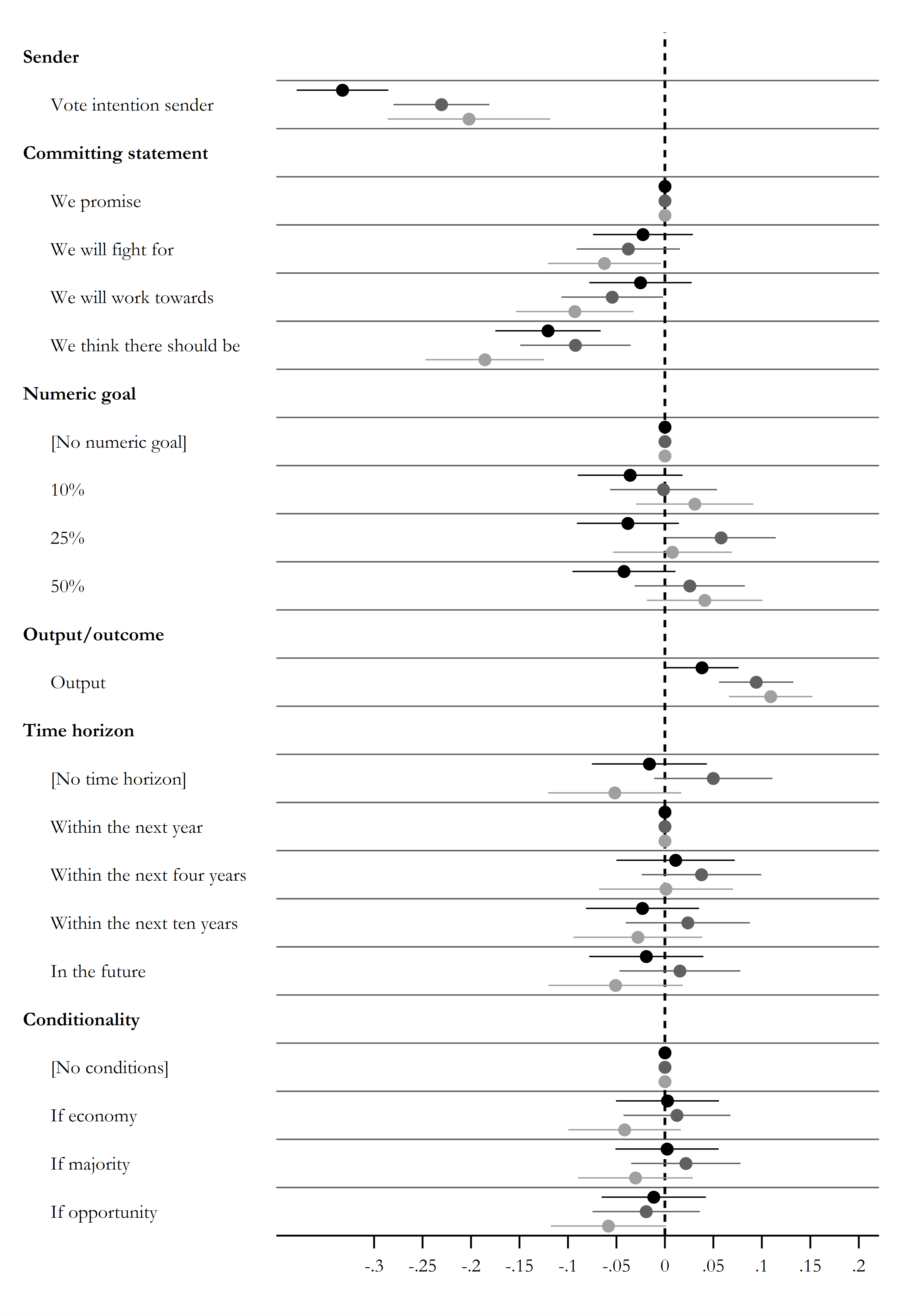
### Appendix F2c: Split-sample, opposition party senders only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering that sender should be held accountable and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

### Appendix F2d: Split-sample, government party senders only



**Note:** Average marginal component effects (AMCE) on the likelihood of considering that sender should be held accountable and 95% confidence intervals. Points without horizontal lines denote the reference category.

Black = United States, dark grey = Britain, and light grey = Denmark.

# References

Abramson, Scott F., Korhan Koçak, and Asya Magazinnik. 2020. “What Do We Learn About Voter Preferences From Conjoint Experiments?” *Working paper.*

Adams, James, Lawrence Ezrow, and Zeynep Somer-Topcu. 2014. “Do voters respond to party manifestos or to a wider information environment? An analysis of mass‐elite linkages on European integration.” *American Journal of Political Science*, 58(4), 967-978.

Aytaç, S. Erdem, Eli G. Rau, and Susan Stokes. 2020. “Beyond opportunity costs: Campaign messages, anger and turnout among the unemployed.” *British Journal of Political Science* 50(4): 1325-1339.

Banducci, Susan, Heiko Giebler, and Sylvia Kritzinger. 2017. “Knowing more from less: How the information environment increases knowledge of party positions.” *British Journal of Political Science*, 47(3), 571-588.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2018. “The number of choice tasks and survey satisficing in conjoint experiments.” *Political Analysis* 26(1): 112–119.

Bisgaard, Martin. 2015. “Bias will find a way: Economic perceptions, attributions of blame, and partisan-motivated reasoning during crisis.” *The Journal of Politics* 77(3): 849–860.

Bisgaard, Martin. 2019. “How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning.” *American Journal of Political Science* 63(4): 824–839.

Bradley, Mark, and Andrew Daly. 1994. “Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data.” *Transportation* 21(2): 167–184.

Dalton, Russell. 2019. *Citizen Politics: Public Opinion and Political Parties in Advanced Industrial Democracies.* Seventh Edition. Washington, DC: CQ Press.

Dupont, Julia C., Evelyn Bytzek, Melanie C. Steffens, and Frank M. Schneider. 2019. “Which Kind of Political Campaign Messages do People Perceive as Election Pledges?” *Electoral Studies* 57: 121–130.

Foa, Roberto Stefan, and Yascha Mounk. 2016. “The Democratic Disconnect.” *Journal of Democracy* 27(3): 5–17.

Foa, Roberto Stefan, and Yascha Mounk. 2017. “The Signs of Deconsolidation.” *Journal of Democracy* 28(1): 5–15.

Funk, Carolyn L. 1996. “The Impact of Scandal on Candidate Evaluations: An Experimental Test of the Role of Candidate Traits.” *Political Behavior* 18(1): 1–24.

Funk, Carolyn L. 1999. “Bringing the Candidate into Models of Candidate Evaluation.” *The Journal of Politics* 61(3): 700–720.

Gillion, Daniel Q., Jonathan M. Ladd, and Marc Meredith. 2020. “Party polarization, ideological sorting and the emergence of the US partisan gender gap.” *British Journal of Political Science* 50(4): 1217-1243.

Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. “Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments.” *Political Analysis* 22(01): 1–30.

Han, Kyung J. 2020. “Beclouding party position as an electoral strategy: Voter polarization, issue priority and position blurring.” *British Journal of Political Science*. 50(2): 653-675.

Hansen, Sune W., Robert Klemmesen, and Søren Serritzlew. 2019. “Losers lose more than winners win: Asymmetrical effects of winning and losing elections.” *European Journal of Political Research* 58(4): 1172-1190.

Hendricks, John Allen, and Robert E. Denton Jr. 2010. *Communicator-in-chief: How Barack Obama used new media technology to win the White House.* Lanham, Maryland: Lexington Books.

Håkansson, Nicklas, and Elin Naurin. 2016. “Promising ever more: An empirical account of Swedish parties’ pledge making during 20 years.” *Party Politics* 22(3): 393–404.

Imai, Kosuke, and Teppei Yamamoto. 2013. “Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments.” *Political Analysis* 21(2): 141–171.

Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago: University of Chicago Press.

Lijphart, Arend. 2012. *Patterns of Democracy*. New Haven: Yale University Press.

Lindgren, Elina. 2017. “Campaigning in Poetry, Governing in Prose: Pre- and Post-Election Effects of Election Pledge Rhetoric.” *Doctoral Dissertation in Political Science.* University of Gothenburg.

Lindgren, Elina. 2018. “Changing Policy With Words: How Persuasive Words in Election Pledges Influence Voters’ Beliefs About Policies.” *Mass Communication and Society* 21(4): 425–449.

Lindgren, Elina, and Elin Naurin. 2017 “Election pledge rhetoric: Selling policy with words.” *International Journal of Communication* 11(22): 2198–2219.

Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. Cambridge: Cambridge University Press.

Mansergh, Lucy, and Robert Thomson. 2007. “Election Pledges, Party Competition, and Policymaking.” *Comparative Politics* 39(3): 311–329.

Mason, Lilliana. 2015. “‘I Disrespectfully Agree’: The Differential Effects of Partisan Sorting on Social and Issue Polarization.” *American Journal of Political Science* 59(1): 128–45.

Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity*. Chicago & London: University of Chicago Press.

Matthieß, Theres. 2020. “Retrospective Pledge Voting: A Comparative Study of the Electoral Consequences of Government Parties’ Pledge Fulfilment.” *European Journal of Political Research*.

Mood, Carina. 2010. “Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It.” *European Sociological Review* 26(1): 67–82.

Müller, Stefan. 2020. “Media Coverage of Campaign Promises Throughout the Electoral Cycle.” *Political Communication*, 1-23.

Naurin, Elin. 2009. “Promising Democracy: Parties, Citizens and Election Promises. *Doctoral Dissertation in Political Science.* University of Gothenburg.

Naurin, Elin. 2011. *Election Promises, Party Behaviour and Voter Perceptions*. Houndsmills: Palgrave Macmillan.

Naurin, Elin. 2014. “Is a Promise a Promise? Election Pledge Fulfilment in Comparative Perspective using Sweden as an Example.” *West European Politics* 37(5): 1046–1064.

Naurin, Elin, Terry J. Royed, and Robert Thomson (eds.). 2019. *Party Mandates and Democracy: Making, Breaking, and Keeping Election Pledges in Twelve Countries*. Ann Arbor, MI: University of Michigan Press.

Naurin, Elin, Stuart Soroka, and Niels Markwat. 2019. “Asymmetric Accountability: An Experimental Investigation of Biases in Evaluations of Governments’ Election Pledges.” *Comparative Political Studies* 52(13–14): 2207–2234.

Royed, Terry J. 1996. “Testing the Mandate Model in Britain and the United States: Evidence from the Reagan and Thatcher Eras.” *British Journal of Political Science* 26(1): 45–80.

Royed, Terry, Elin Naurin, and Robert Thomson. 2019. “A systematic research design for studying the fulfillment of election pledges.” In Elin Naurin, Terry J. Royed, and Robert Thomson (eds.) *Party Mandates and Democracy: Making, Breaking, and Keeping Election Pledges in Twelve Countries*. Ann Arbor, MI: University of Michigan Press.

Schedler, Andreas. 1998. “The normative force of electoral promises.” *Journal of Theoretical Politics*, 10(2): 191-214.

Spoon, Jae-Jae, and Klüver, Heike. 2019. “Party convergence and vote switching: Explaining mainstream party decline across Europe.” *European Journal of Political Research* 58(4): 1021-1042.

Thomson, Robert. 2001. “The Programme to Policy Linkage: The Fulfilment of Election Pledges on Socio-Economic Policy in The Netherlands, 1986–1998.” *European Journal of Political Research* 40(2): 171–197.

Thomson, Robert. 2011. “Citizens’ evaluations of the fulfillment of election pledges: Evidence from Ireland.” *The Journal of Politics* 73(1): 187–201.

Thomson, Robert, Terry Royed, Elin Naurin, Joaquín Artés, Rory Costello, Laurenz Ennser-Jedenastik, Mark Ferguson, Petia Kostadinova, Catherine Moury, François Pétry, and Katrin Praprotnik. 2017. “The Fulfillment of Parties’ Election Pledges: A Comparative Study on the Impact of Power Sharing.” *American Journal of Political Science* 61(3): 527–542.

Tilley, James, and Sara B. Hobolt. 2011. “Is the government to blame? An experimental test of how partisanship shapes perceptions of performance and responsibility.” *The Journal of Politics* 73(2): 316–330.

Schattschneider, Elmer E. 1942. *Party Government*. New York: Rinehart and Winston.

1. Question wording: How much do you agree with the following statement: This is a high priority for [sender]? [↑](#footnote-ref-1)
2. Question wording: How much do you feel you can trust [sender]? [↑](#footnote-ref-2)
3. Question wording: Is this something [sender] can do something about if [sender] really wants to? [↑](#footnote-ref-3)
4. Question wording: How much do you agree with the following statement: In reality, it is impossible for people like me to check whether [sender] actually implements this? [↑](#footnote-ref-4)