

# **The Influence of Gender Stereotypes on Politicians' Behaviour and Voter Attitudes**

**Lotte Hargrave**

**PhD in Political Science,  
Department of Political Science,  
University College London (UCL)**

**October, 2022**

I, Lotte Hargrave, confirm that the work presented in this thesis is solely my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

# Abstract

Do politicians behave in accordance with gender stereotypes? Does the pressure to do so shift by context and over time? Do voters uphold these stereotypes when they perceive and evaluate politicians' behaviour? Focusing on elites and voters in the United Kingdom, this thesis addresses these questions in three papers. The first paper analyses whether politicians express behavioural styles that are consistent with stereotypes, and how the pressure to do so may have diminished over time. I describe novel quantitative text analysis approaches to measure styles in parliamentary debates between 1997 and 2019 and show that women's styles have changed substantially over time, as they have increasingly adopted styles congruent with "masculine" stereotypes. The second paper investigates whether there are gender differences in the sets of issues that politicians raise, whether women do more to raise women's concerns and experiences across the policy process more broadly, and how the incentive to do so changes with increased political experience. Using quantitative text analysis techniques to measure the issues politicians raise, I find that, among junior politicians, women talk significantly more about stereotypically "feminine" issue areas, but also that this gender gap decreases markedly with increased seniority. However, women continue to refer to women's experiences in a wide range of issue areas throughout their careers. The third paper evaluates whether the styles politicians use influences how voters evaluate them, and whether this matters more for women than it does for men. In a novel survey experiment, I manipulate politician gender and argument style and find that style usage has important consequences for how voters evaluate politicians, but that the effects of style are not conditional on politician gender. Taken together, this thesis provides important theoretical arguments and empirical evidence concerning the dynamic validity of stereotypes in informing elite and voter behaviour in the UK.

# Impact statement

My doctoral research sits at the intersection of a diverse range of political science sub-fields, and therefore advances academia in the areas of gender and stereotyping, political communication, behaviour, and legislative studies. First, I translate the conceptual definitions of gender stereotypes that emerge in the social psychology literature into theoretically grounded observable behavioural indicators that reflect the ideas of communality and agency. Second, I take a dynamic approach to studying the accuracy of stereotypes for informing politicians' behaviour, by measuring how the issues politicians raise in parliamentary debate, and the style in which they deliver these arguments, may shift by context and over time. In doing so, I show that the accuracy of stereotypes is far from fixed and highlight instead the dynamic influence they exert over elite behaviour. Third, I provide a methodological contribution to the measurement of styles and issues in legislative settings that will be of direct value to researchers studying politicians' speechmaking behaviour. I introduce a new measurement approach that combines dictionaries with a word-embedding model to detect my concepts of interest in the specific context of parliamentary debate. Fourth, I present a theoretical argument and design a survey experiment that attempts to unpack some of the mechanisms which underpin gender bias in voters' evaluations. This approach provides new insights on gender bias in voting behaviour and attitudes.

Understanding what shapes the communication strategies that politicians deploy, and how voters engage with this behaviour has policy implications in four areas. First, I provide evidence that may be of direct value to think tanks, research organisations, and activist groups, such as the Fawcett Society and 50:50 Parliament, who are engaged in the study and promotion of women into office. When running, women are now no less likely to win than men, yet barriers remain in their likelihood to stand. My work shows that not only may restrictions on women's behaviour be weaker now than in the past, but also that voters do not punish women any more than men for counter-stereotypical behaviour. Second, I show that as women progress through their parliamentary careers, they increasingly raise women's perspectives in a diversity of issue areas, thus ensuring that the interests of women voters are raised in debates across the policy spectrum. It is therefore important to not only increase women's presence in politics, but also to ensure they reach sufficiently senior levels within it. Third, I provide new information on the kinds of communication strategies that voters value in politicians. This may be of use to political parties and organisations such as the Commonwealth Parliamentary Association, which provides training and expertise, or the Global Partners Governance which provides advice to politicians, political organisations, and parliaments. Fourth, I provide evidence that may be of interest to practitioners and proponents of parliamentary reform, who are dedicated to designing institutions that offer effective working environments for men and women. The attention I bring to gendered behaviour may provide pointers for ongoing efforts in achieving a "Gender Sensitive" parliament.

# Acknowledgements

The words “get the thing out of the door and move on with your life” have been written on a post-it note that has been stuck to my monitor for the past year. But now the time to move on has finally come, so too has the time to reflect upon, and thank, the army of people who have helped me get to this point and infinitely improved my journey here.

First of all, thank you to my formal supervision team, Meg Russell, Jack Blumenau, and Jennifer Hudson, who, in some cases, have seen me through from even before the beginning of the PhD process and, in all cases, have seen me through to the end of it. Thank you Meg: it was due to your encouragement, and early belief in my ability to do so, that I decided to pursue a PhD at UCL in the first place. While the content of this thesis has evolved a little from what we discussed in those earlier conversations, your guidance, perceptive feedback, eye for detail, and sharing of your boundless knowledge of the inner workings of the UK parliament have been vital in shaping the quality of it. Thank you Jennifer: my time as a PhD student has coincided with your tenure as Head of Department, and despite the scale of this commitment, you have never failed to be generous with your time, support, kindness, feedback, and reassurance that the downs will always be accompanied by the ups. When my first paper was accepted you took me out for cocktails and sent me an abstract mug to mark its publication, thank you for this occasion, and for the many others, when you have taken the time to ensure that I have always felt properly celebrated. Thank you Jack: for acting as a methods tutor, grammatical advisor, informal therapist, regular provider of caffeinated beverages, role model, and friend. When you agreed to join my supervision team, it was initially under the provision of offering occasional “methods help”, and I can only apologise for quite how much more than simply methods help your role has ended up entailing. Whilst (sorry) your help, guidance, and advice have been instrumental in greatly improving the substantive and methodological content and rigour of this thesis overall, co-authoring a paper with you has taught me a lot, improved my skills and abilities as a researcher, and made what could have been a very isolating PhD experience a lot less lonely. Above all else, there has been more than one occasion over the past few years where I sincerely believe I would have given up on this PhD had it not been for your never-ending support, encouragement, kindness, patience, belief in me, and constant reassurance that everything is in fact okay. I really could not have done it without you, nor would I have wanted to do so.

Not only have I been fortunate enough with my formal supervision team, but I have been the grateful recipient of help from an army of informal supervisors. For reading essentially everything I have ever written, teaching me not to critique the approaches of others if I don't have a better solution, and for acting as a role model for how to be the very best team player, I thank Lucy Barnes. For the football matches, beers, unwavering belief in me, and for being the best methods teaching companion

---

I could have asked for, I thank Judith Spirig. For the baked goods, rescuing me from the basement all those years ago, provision of ever-honest advice on how to navigate the process, and for being an exceptional colleague and friend, I thank Paolo Morini. For letting me crash several of his lectures, making my Upgrade Viva one of the nicest experiences of my PhD, and for leading the CPEB group, I thank Tom O'Grady. For her patience and help when I first started teaching, constant encouragement, and unwavering kindness, I thank Katya Tertytchnaya. For teaching me it's okay to say you don't understand and for always making me laugh, I thank Tim Hicks. For the wisdom on experimental designs and for the reminder that I am young and have plenty of time to change my mind, fail, and try again, I thank Chris Wratil. For the early methodological advice and for teaching me that one piece of research cannot, and should not, do everything, I thank Ben Lauderdale. For doing pretty much all of the above, and so, so much more than I have sufficient space to truly do justice to, I thank Becky McKee and Jess Smith. Thank you for your friendships, being there before, throughout, and I am certain after, keeping me sane during the darker times, all of the fun and laughter in the lighter times, the many research discussions, feedback on my work, glasses of wine, holding me accountable, and reminding me that work really isn't everything.

The cliché that “a PhD can be a lonely experience” has proven not to be true in my case thanks to being surrounded by wonderful PhD colleagues and friends who have helped drag me through it. To the women in my cohort, Alice, Angelica, Jennifer, and Sigrid, we've been through a lot together, but somehow, we've all made it through, and I feel exceptionally lucky to have shared the journey with you all. A special mention to Alice, for the bracing swims and for always managing to make me feel better about myself. And to Sigrid, my “4th supervisor”, for the constant R help, putting up with my unrelenting need to chat, the morning gossiping, and for being a true rock. Although I am sure she would not like me to admit it, she has been a permanent source of reassurance, support, and encouragement that I could not have made it through without. For the teaching companionship, regular reminder that I needed a better “grand theory”, and for always making turning up to work a far brighter prospect, I thank Kit. For ensuring lunchtime conversation agendas were always full and for guiding me through my first PhD emotional crisis (and many more since then), I thank Kasim. For being my trusted discussant, provider of jokes that always made me laugh (although quite often at and not with him), and for being my cherished political behaviour companion throughout in this otherwise conflict-dominated department, I thank Markus. For kindly proofreading parts of this thesis and for allowing me to bounce various research ideas off of him throughout the past year, I thank Michael. For encouraging me to always be ambitious in approach, the many pep talks, and for the shared love of watching dogs play in the park, I thank Philipp. For all of their sage advice, kindness, encouragement, regular sharing of exasperation at the difficulty of the process, and for being exceptional former PhD siblings, I thank Ruxandra and Roberta. To those in the years above mine, thank you for your generous guidance on how to best make it through, showing me that it is actually possible to do so, and for setting the standard for how PhDs should treat one another. To those in the years below mine, my apologies for the imparting of all of the “advice”, I can only hope it ends up being useful, and thank you for filling me with confidence that this PhD community will continue to be a kind and supportive one.

To those in the CPEB research cluster, thank you for your constant willingness to read and comment on my work and for acting as role models for how to provide feedback in a way that is helpful,

perceptive, thoughtful, and, importantly, kind. To those at the Constitution Unit, thank you for the virtual teas, endless supply of cake, constant reminder of just how much I still have to learn about British politics, and for my time as a researcher. For acting as a co-author, my cheerleader, and for being with me on my first journey into the literature on gendered styles, I thank Tone. For serving as my PhD mother figure, I thank Helen. For all of the wine and general good cheer, I thank Ben. To many others in the UCL Department of Political Science, thank you for making this a great place to have spent the past six years of my life.

Finally, my thanks go to the people who have supported me from long before the PhD process began, and who I know will continue to support me long after the end of it. To my friends, who have always been there to hype me up and, quite rightly, tell me when it was time to shut up. To my dogs, Sophie and Bella, who obviously have no idea what a PhD is, but who have had to dole out plenty of cuddles to help me get through it. To my sister, Anna, who unfortunately had to bear the brunt of living with me through a large chunk of the pandemic element of my PhD. Your love, patience, support, tolerance for my worrying, and ever-flowing provision of Sauvignon Blanc made all the difference. To my parents, Sue and Peter, for their unwavering support, generosity, love, and never-ending belief that I could and would make it to the end of this PhD and anything else I put my mind to.

# Contents

<b>Abstract</b>	<b>3</b>
<b>Impact statement</b>	<b>4</b>
<b>Acknowledgements</b>	<b>5</b>
<b>1 Introduction</b>	<b>14</b>
1.1 Gender roles and stereotypes . . . . .	17
1.2 Substantive approach and contributions . . . . .	19
1.3 Methodological approach and contributions . . . . .	27
1.4 Implications for gender and politics . . . . .	30
<b>2 Literature Review</b>	<b>34</b>
2.1 The formation of gender stereotypes . . . . .	35
2.2 The application of gender roles to politics . . . . .	39
2.3 Style-based behavioural differences . . . . .	41
2.4 Policy-based behavioural differences . . . . .	45
2.5 Stereotypes and voters' perceptions and evaluations . . . . .	48
2.6 Conclusion . . . . .	54
<b>3 No Longer Conforming to Stereotypes? Gender, Political Style, and Parliamentary Debate in the UK</b>	<b>57</b>
Abstract . . . . .	57
3.1 Introduction . . . . .	58
3.2 Gender, stereotypes, and debating style . . . . .	60
3.3 Data and methodology . . . . .	66
3.4 Modelling political style . . . . .	72
3.5 Results . . . . .	75
3.6 Conclusion . . . . .	81
<b>4 Earning Their Stripes? Gender, Political Experience, and Policy Prioritisation in Parliamentary Debate</b>	<b>83</b>
Abstract . . . . .	83
4.1 Introduction . . . . .	84
4.2 Descriptive and substantive representation: Evidence and explanations . . . . .	87
4.3 Data and methodology . . . . .	94
4.4 Empirical strategy . . . . .	96
4.5 Results . . . . .	98
4.6 Amplifying women's experiences across the policy process . . . . .	103
4.7 Conclusion . . . . .	108
<b>5 A Double Standard? Gender Bias in Voters' Perceptions of Political Arguments</b>	<b>111</b>



Abstract . . . . .	111
5.1 Introduction . . . . .	112
5.2 Gender, stereotypes, and voter backlash . . . . .	115
5.3 Experimental design . . . . .	119
5.4 Methodology . . . . .	127
5.5 Results . . . . .	129
5.6 Conclusion . . . . .	137
<b>6 Conclusion</b>	<b>142</b>
6.1 Key findings and substantive contributions . . . . .	142
6.2 Implications for politics . . . . .	150
6.3 Generalisability, limitations, and avenues for future work . . . . .	153
<b>A Appendix A – No Longer Conforming to Stereotypes? Gender, Political Style, and Parliamentary Debate in the UK</b>	<b>159</b>
A.1 Word-embedding-based dictionaries . . . . .	159
A.2 Validation tests . . . . .	163
A.3 Controlling for individual-level covariates . . . . .	170
A.4 Style use and debate-type . . . . .	174
A.5 Within-MP and replacement effects . . . . .	178
A.6 Topic-based confounding . . . . .	182
A.7 Style use and debate participation . . . . .	188
<b>B Appendix B – Earning Their Stripes? Gender, Political Experience, and Policy Prioritisation in Parliamentary Debate</b>	<b>191</b>
B.1 Seed words . . . . .	191
B.2 Validation tests . . . . .	193
B.3 Measure correlation . . . . .	213
B.4 Structural topic model-based seniority analysis . . . . .	214
B.5 Robustness checks . . . . .	217
B.6 Controlling for individual-level covariates . . . . .	224
B.7 Raising women’s perspectives across issue areas . . . . .	226
B.8 Gender differences in policy areas over time . . . . .	228
<b>C Appendix C – A Double Standard? Gender Bias in Voters’ Perceptions of Political Arguments</b>	<b>232</b>
C.1 Experiment pre-test survey . . . . .	232
C.2 Sample recruitment . . . . .	238
C.3 Treatment texts . . . . .	239
C.4 Full models . . . . .	244
C.5 Multiple comparisons correction . . . . .	250
C.6 Pooled models . . . . .	253
C.7 MP gender and policy area . . . . .	258
C.8 Heterogeneous effects by voter gender . . . . .	261
C.9 Power analysis . . . . .	265

# List of Figures

1.1	Number of women MPs by party, 1992-2019 . . . . .	21
1.2	Percentage of respondents who agree or disagree with the statement “a man’s job is to earn the money; a woman’s job is to look after the home and family”, 1984-2017 . . . . .	22
3.1	Human validation task prompt . . . . .	71
3.2	Gender differences in style over time . . . . .	76
3.3	Average style use for men and women over time . . . . .	79
4.1	Gender differences in issue areas . . . . .	99
4.2	Gender gaps in issues over parliamentary tenure . . . . .	100
4.3	Average topic use for men and women over parliamentary tenure . . . . .	102
4.4	Effect of time in parliament on gender gaps in raising women’s perspectives . . . . .	104
4.5	Gender differences in raising women’s perspectives across debates on different topics . . . . .	105
4.6	<b>Effect of years in parliament on topic discussion for men and women:</b> The figure shows a two-way fixed effect specification with MP and parliamentary term fixed effects. . . . .	108
5.1	Example experiment prompt (Evidence, statistics, transport, woman) . . . . .	126
5.2	Unconditional relationship between control and treatment style usage and style perceptions, MP likeability, and MP competence evaluations . . . . .	130
5.3	Conditional relationship between MP gender, style treatment group, style perceptions, MP likeability, and MP competence evaluations . . . . .	132
5.4	Conditional relationship between voter gender, style treatment status, style perceptions, MP likeability, and MP competence evaluations . . . . .	136
5.5	Conditional relationship between MP gender, voter gender, style treatment status, style perceptions, MP likeability, and MP competence evaluations . . . . .	138
A.1	Aggression seed dictionary . . . . .	160
A.2	Human narrative seed dictionary . . . . .	160
A.3	Gender differences in style over time controlling for individual-level confounders . . . . .	171
A.4	Individual-level covariate effects . . . . .	173
A.5	Debate type models . . . . .	175
A.6	Style type average by debate type . . . . .	177
A.7	Within-MP and replacement over time change by gender . . . . .	180
A.8	<b>Topical-confounding:</b> The figure shows the relationship between the gender gap in the use of a given style on a given topic (x-axis), and the change in the prevalence of a given topic over time (y-axis). . . . .	185
A.9	<b>Topical-confounding, varying <math>K</math>:</b> On the y-axis, the figure summarises the linear relationship between the gender gap in the use of a given style on a given topic ( $\delta_k^s$ ), and the change in the prevalence of a given topic over time ( $\gamma_k$ ). The x-axis measures the number of CTM topics, $K$ , used to estimate these relationships. . . . .	187

A.10	<b>Participation as a function of average style use, by parliamentary term:</b> The figure illustrates the average marginal effect of a one standard deviation increase in the average style use on the number of times an MP speaks in a given parliamentary term. . . . .	189
B.1	Average topic use by government, opposition, and committee positions . . . . .	205
B.2	Average topic use in debates on the same topics . . . . .	212
B.3	Correlation between issue areas . . . . .	213
B.4	<b>Effect of parliamentary experience on issue prevalence:</b> The figure shows the relationship between years in parliament and the CTM topics, <i>K</i> . . . . .	216
B.5	Individual-level covariate effects . . . . .	225
B.6	Gender differences in issues over time . . . . .	229
B.7	Average topic use for men and women over time . . . . .	231
C.1	Example of pre-testing task (Aggression, treatment, health) . . . . .	233
C.2	Distribution of responses across treatment group status for emotion . . . . .	234
C.3	Distribution of responses across treatment group status for aggression . . . . .	235
C.4	Distribution of responses across treatment group status for evidence . . . . .	236
C.5	Power analysis . . . . .	266

# List of Tables

3.1	Political styles . . . . .	63
3.2	Correlation between text-based measures and human judgements . . . . .	71
5.1	Style definitions . . . . .	116
5.2	Vignette examples . . . . .	122
A.1	Word-level validation . . . . .	166
A.2	Top sentences for Affect, Positive Emotion, and Human Narrative . . . . .	168
A.3	Top sentences for Aggression, Fact, and Negative Emotion . . . . .	169
B.1	Seed words . . . . .	191
B.2	Word-level validation . . . . .	197
B.3	Top sentences for Agriculture, Children and Family, and Crime and Policing . . . . .	199
B.4	Top sentences for Defence, Education, and Environment . . . . .	200
B.5	Top sentences for Finance and Economy, Health, and Trade . . . . .	201
B.6	Top sentences for Transport and Welfare . . . . .	202
B.7	Top debates for Agriculture, Children and Family, Crime and Policing . . . . .	207
B.8	Top debates for Defence, Education, and Environment . . . . .	208
B.9	Top debates for Finance and Economy, Health, and Trade . . . . .	209
B.10	Top debates for Transportation and Welfare . . . . .	210
B.11	Relationship between MP gender and years in parliament on agriculture topic . . . . .	219
B.12	Relationship between MP gender and years in parliament on children and family topic . . . . .	219
B.13	Relationship between MP gender and years in parliament on crime and policing topic . . . . .	220
B.14	Relationship between MP gender and years in parliament on defence topic . . . . .	220
B.15	Relationship between MP gender and years in parliament on education topic . . . . .	221
B.16	Relationship between MP gender and years in parliament on environment topic . . . . .	221
B.17	Relationship between MP gender and years in parliament on economy topic . . . . .	221
B.18	Relationship between MP gender and years in parliament on health topic . . . . .	222
B.19	Relationship between MP gender and years in parliament on trade topic . . . . .	222
B.20	Relationship between MP gender and years in parliament on transport topic . . . . .	222
B.21	Relationship between MP gender and years in parliament on welfare topic . . . . .	223
B.22	Relationship between MP gender and years in parliament on references to women . . . . .	223
B.23	Relationship between MP gender and years in parliament on amplifying women’s per- spectives across debate types . . . . .	227
B.24	Relationship between MP gender and years in parliament on amplifying women’s per- spectives across debate types . . . . .	227
C.1	Pre-testing results . . . . .	232
C.2	Gendered names for experimental design . . . . .	239
C.3	Treatment texts . . . . .	240
C.4	Relationship between style usage and style perceptions . . . . .	244

C.5	Relationship between style usage and likeability evaluations . . . . .	244
C.6	Relationship between style usage and competence evaluations . . . . .	245
C.7	Relationship between MP gender and MP likeability evaluations . . . . .	247
C.8	Relationship between MP gender and MP competence evaluations . . . . .	248
C.9	Relationship between MP gender and style perceptions . . . . .	249
C.10	Comparison between unadjusted and Benjamini-Hochberg adjusted p-values . . . . .	252
C.11	Relationship between female stereotype congruent styles, likeability, and competence evaluations . . . . .	253
C.12	Relationship between MP gender, style usage, likeability, and competence evaluations . . . . .	256
C.13	Relationship between MP gender, policy area, likeability, and competence evaluations . . . . .	260
C.14	Relationship between voter gender and likeability evaluations, competence evaluations, and emotion perceptions . . . . .	262
C.15	Relationship between voter gender and likeability evaluations, competence evaluations, and aggression perceptions . . . . .	262
C.16	Relationship between voter gender and likeability evaluations, competence evaluations, and evidence perceptions . . . . .	262
C.17	Relationship between MP gender, voter gender, and likability, competence, and perceived emotion for the emotion style . . . . .	264
C.18	Relationship between MP gender, voter gender, and likability, competence, and perceived aggression for the aggression style . . . . .	264
C.19	Relationship between MP gender, voter gender, and likability, competence, and perceived evidence for the evidence style . . . . .	264

Word count: 59,134

# Chapter 1

## Introduction

Gender stereotypes are said to wield significant influence over the behaviour and experiences of women in public and political life. It is not difficult to summon examples from electoral campaigns and political office of women being scrutinised by their colleagues, the media, and voters for how they behave. In the 2008 US Presidential Election, Michelle Obama was condemned as an “angry black woman” and, allegedly as a result of this characterisation, went on to soften her image and speeches in order to be seen as more “likeable”. In the 2016 US Presidential Election, Hillary Clinton was described as an “unlikeable”, “nasty”, and “shrill” woman, and the wider campaign context was coloured by slogans such as “Trump that bitch”, “Shrillary”, and “Killary”. Examples of the scrutiny of women’s behaviour persisted into the 2020 US Presidential Election, where the supposed “anger and antagonism” of prominent women candidates Kamala Harris and Elizabeth Warren was a persistent feature of the news cycle. Media commentators argued that anger is “not a good look” for them and their opponents criticised them for being “angry”, “nasty”, and “vicious”.

Such examples extend beyond the US. In the UK, former UK Prime Minister Theresa May was characterised as the “unemotional”, “inhuman and uncaring” and “cold” “Maybot” because her leadership style contrasted sharply with stereotypes about the supposed emotional sensitivity of women. Inside the legislature, women politicians in the UK have been accused of being “extremely frustrated”, told to “watch their tone”, and to “calm down, dear”. In New Zealand, Prime Minister Jacinda Ardern was told to “zip it, sweetie” during a parliamentary debate. Former Australian Prime Minister Julia Gillard has spoken openly about the regular misogynistic and sexist abuse she received during her time as the country’s first woman leader, where she was called “hysterical” and “brittle”, and endured sexist protests outside of the Australian Parliament, which were joined by prominent politicians, demanding parliament should “ditch the witch”.<sup>1</sup>

---

<sup>1</sup>Throughout this thesis, I opt to use “man”/ “men” and “woman”/ “women” as both adjectives and nouns. I do this for two

---

These examples illustrate the scrutiny of, and behaviour towards, women in political life when they fail to adhere to gender stereotypical behavioural expectations. Widely held stereotypical descriptions state that women are expected to be nice, likeable, warm, kind, compassionate, communal, nurturing, people-orientated, empathetic, and concerned for the welfare of others. By contrast, stereotypical behavioural expectations traditionally associated with men orientate towards assertiveness, aggression, confidence, competence, ambition, self-sufficiency, logic, self-confidence, and goal attainment (Diekman and Eagly, 2000; Eagly and Karau, 2002; Fiske and Stevens, 1993; Huddy and Terkildsen, 1993a; Schneider and Bos, 2014, 2019). The experiences of Michelle Obama, Hillary Clinton, Kamala Harris, Elizabeth Warren, Theresa May, Nadine Dorries, Rosena Allin-Khan, Angela Eagle, Jacinda Ardern, Julia Gillard, and many other women in political life exemplify the treatment of women when they fail to live up to the expected behavioural standards of their gender or are perceived to be doing so by others. How might these dynamics operate for men in political office? Or, put simply, would an angry, tough, or emotional man have been branded shrill, nasty, robotic, or hysterical? While, of course, it is impossible to say in any given instance, existing evidence suggests that this is unlikely to be the case (Boussalis et al., 2021; Gleason, 2020; Okimoto and Brescoll, 2010).

Do politicians behave in accordance with traditional gender stereotypes? Does the pressure to do so shift by context and over time? Do voters uphold these stereotypes when they perceive and evaluate politicians' behaviour? In this thesis, in the context of the United Kingdom (UK), I set about answering these three questions. The critiques of women's behaviour described above are each underpinned by a set of gendered stereotypical expectations that are the product of the positions that men and women have historically occupied in politics specifically and society more broadly (Eagly and Karau, 2002). However, the content of such stereotypes is not necessarily fixed but should instead be considered *dynamic* concepts (Diekman and Eagly, 2000; Eagly and Wood, 2012; Eagly et al., 2020). The stereotypic characteristics of a group are related to the activities associated with their traditional roles and divisions of power in society. Therefore, as the distribution of positions and power in society and politics that men and women occupy have shifted (Bianchi et al., 2014; Blau, Brummund and Liu, 2013; Diekman and Goodfriend, 2006; Sayer, Bianchi and Robinson, 2004), the stereotypical characteristics associated with their genders may have shifted too. As such, stereotypes may inform both how men and women behave and the standards to which they are held by others, however, the content of these

---

reasons. First, the central questions I assess focus on how traditional gender stereotypes inform the behaviour of politicians in office and how voters engage with this behaviour. As these questions relate to the prominence of societal and political *gender*-based roles and not biological *sex*-based roles, it is appropriate to use "men"/ "women" and not "male"/ "female". Second, although from a traditional grammatical perspective "male"/ "female" act as both nouns and adjectives and "men"/ "women" only as nouns, there has been a move in recent years to accept the use of "men" and "women" as adjectives in order to adapt to inclusive language use.

stereotypes may shift over time as the distribution and power of men and women in different social roles have also shifted. Many of the criticisms and behavioural expectations described above may be rooted in a potentially anachronistic and outdated view of the extent to which men and women both conform to and are held to the standards of traditional “masculine” and “feminine” behaviours.

At present, we do not have clear evidence on whether contemporary politicians fulfil these behavioural stereotypes, whether such stereotypes inform voter behaviour, nor how this may have changed over time. In this thesis, I make an important contribution to knowledge by theorising and testing empirically whether politicians behave in accordance with traditional stereotypes with respect to the policy areas they champion in parliamentary debate, the stylistic manner in which they deliver these arguments, and how the pressure to do so may shift both over time and throughout the course of their careers. Further, I consider whether stereotypes inform how voters engage with politicians’ behaviour. I assess, first, whether voters perceive men and women politicians’ behaviour differently based on stereotypical expectations, second, whether politicians incur backlash when violating gender-based norms, and third, whether differential perception of behaviour may itself explain this backlash. I assess both whether politicians actually behave in a way that is consistent with gender stereotypes, but also whether voters simply perceive there to be differences based on stereotypical assumptions. Taken together, the three papers which form the core of this thesis aim to leverage creative and theoretically informed quantitative designs to provide insights on the *dynamic* validity of stereotypes in informing contemporary elite and voter behaviour in the UK.

Contrary to past work, my results suggest that not only are the descriptive validity of prominent gender stereotypes for how elites behave considerably lower than they were in the past, but also that voters do not punish women in the way that prominent theories of gender stereotyping predict. Documenting when behaviour and attitudes run counter to expectations is of central normative importance because it may help to both undermine and potentially update outdated stereotypical prescriptions, and thereby further reduce voter bias against women. However, as the opening paragraph demonstrates, sexism persists as a serious issue in politics and public life. I make an important contribution by shedding light on several areas where gendered bias may have declined in recent years. However, these findings should be considered within the wider context of women’s experiences in politics. The findings should motivate researchers to refocus their attention on understanding how sexism, stereotyping, and gendered inequality shape women’s experiences in political life more broadly.

In the remainder of this introductory chapter, I outline the theoretical literature on gender roles and stereotype formation, as well as how stereotypes may induce differences in politicians’ behaviour.



Next, I introduce the UK context in which I situate this study and turn to the substantive and methodological approaches and contributions that each paper provides. Finally, I conclude by discussing the wider implications of these findings for women's experiences in politics.

## 1.1 Gender roles and stereotypes

Throughout this thesis, I adopt the general framework of gender role theory to understand the influence of stereotypes in politics (Eagly, 2013; Eagly and Karau, 2002; Schneider and Bos, 2019). Gender role theory suggests that stereotypes are formed initially through the traditional division in societal roles between men and women. These gender divisions are associated with expectations for how individuals should think, feel, and behave. These expectations are reflected and reinforced in many ways: through educational experiences, social interactions, workplace cultures, and numerous other cumulative experiences. Over time, men and women come to *internalise* the traits that align with their gender roles, which are then further reinforced by others who reward conformity to expected behaviours and punish violations. Further, individuals internalise the expectation that men and women *should* behave in ways that are consistent with their gender roles. Or, put differently, these descriptive stereotypes, which describe how men and women *typically* behave, lead to the formation of prescriptive stereotypes, which outline instead how men and women *should* behave (Gill, 2004; Heilman, 2001). As such, the idea that women are generally warm and caring becomes a prescription that women *should* be warm and caring. The violation of these behavioural prescriptions can result in social sanctions imposed by others which come to incentivise conformity with these gender-based expectations (Brescoll and Uhlmann, 2008). Gender roles then become self-reinforcing, as societal beliefs about the typical behavioural differences lead to ever-more entrenched patterns of gendered behaviour. The consequence, therefore, is that women come to behave in feminine, or “communal”, ways that are associated with positivity, warmth, compassion, kindness, and caring for others, and men come to behave in masculine, or “agentic”, ways that are associated with strength, assertiveness, logic, and independence (Bakan, 1966; Best and Williams, 1990; Broverman et al., 1972; Eagly and Steffen, 1984; Rucker, Galinsky and Magee, 2018; Sczesny, Nater and Eagly, 2019; Spence, Helmreich and Stapp, 1974).

In the political realm, numerous studies have sought to assess the influence gender stereotypes wield on both politicians' behaviour and the ways in which voters engage with this behaviour (Bauer, 2020b; Bernhard, 2022; Boussalis et al., 2021; de Geus et al., 2021; Gleason, 2020; Huddy and Terkildsen, 1993a; Krupnikov and Bauer, 2014; Saha and Weeks, 2022; Teele, Kalla and Rosenbluth, 2018). Stereo-

types are said to inform both the styles that politicians express and the set of policy responsibilities they are deemed suited to and on which they are considered competent (Huddy and Terkildsen, 1993a; Kahn, 1996; Lawless, 2004; McDermott, 1997). Consistent with the idea that women are, and should be, “communal”, they are expected to express styles that are marked by higher levels of emotionality, positivity, empathy, warmth, and human interest (Alexander and Andersen, 1993; Blankenship and Robson, 1995; Dahlerup, 1988; Karpowitz and Mendelberg, 2014). Further, this may also feed into the expectation that women have strengths in “compassionate” issue areas that are concerned with caring for others, such as children and family, poverty, welfare, education, and health (Eagly, 2013; Herrnson, Lay and Stokes, 2003; Huddy, Cassese and Lizotte, 2008; Sanbonmatsu, 2002; Shapiro, 1986). Similarly, consistent with the idea that men are, and should be, “agentic”, they are expected to express styles marked by higher levels of assertiveness, aggression, logic, and confidence (Huddy and Terkildsen, 1993a; Jamieson, 1995; Mendelberg, Karpowitz and Oliphant, 2014), and gravitate towards associated policy areas such as the economy, finance, defence, or employment (Bäck, Debus and Müller, 2014; Escobar-Lemmon and Taylor-Robinson, 2005; Krook and O’Brien, 2012).

Observational work assessing *elite behaviour* broadly supports these perspectives, and finds that women tend to be more emotional (Dietrich, Hayes and O’Brien, 2019; Gleason, 2020; Gleason, Jones and McBean, 2019; Jones, 2016), positive (Boussalis et al., 2021; Yu, 2013), less aggressive or dominant (Grey, 2002; Karpowitz and Mendelberg, 2014; Mendelberg, Karpowitz and Goedert, 2013; Kathlene, 1994; Shaw, 2000), and tend to draw more on human narrative and experiences (Bligh and Kohles, 2008; Childs, 2004b; Hargrave and Langengen, 2021; Karvonen, Djupsund and Carlson, 1995; Mattei, 1998). Further, in electoral campaigns (Ennser-Jedenastik, 2017) and in the legislature (Bäck and Debus, 2019) women tend to talk more about social policy issues than men and take on legislative roles that reflect this division (Blumenau, 2021c; Coffé, Bolzendahl and Schnellecke, 2019; Goodwin, Holden Bates and Mckay, 2021).

Empirical assessments of *voter behaviour* suggest that voters may punish politicians who contradict gender stereotypical styles (Bauer, 2015a; Boussalis et al., 2021; Cassese and Holman, 2018), and ascribe women more qualifications and competence in “feminine” policy areas (Herrnson, Lay and Stokes, 2003; Huddy and Terkildsen, 1993a; Sanbonmatsu and Dolan, 2009), while punishing women who do not advance “women’s issues” (Cassese and Holman, 2018). Several other studies have suggested that stereotypes *overall* may not inform voters’ evaluations, but that stereotype application depends on the geographic (Saha and Weeks, 2022) or electoral context (Anzia and Bernhard, 2022), or on the type of voter (Bauer, 2015b; Bauer, Kalmoe and Russell, 2022). Taken together, these accounts

suggest that, prior to entry into politics, politicians may internalise expectations from their gender roles, and the idea that they may incur punishments from voters if they fail to live up to these expectations. As such, legislators may engage in gender-role consistent behaviours and voters may hold them to these standards.

Gender role theory informs each of the papers in this thesis in several key ways. First, gender role theory provide conceptual definitions of gender stereotypical behaviour which, while not always precisely defined, are broadly described as communal behaviours which are associated with women, and agentic behaviours which are associated with men. I translate these conceptual definitions into theoretically grounded behavioural indicators that allow me to measure whether politicians conform to stereotypes both with respect to the styles they express, and the substantive issues which they champion. Second, gender role theorists contend that stereotypes are dynamic concepts and that the content of stereotypes associated with men and women may change over time as the distribution of gendered societal positions have changed. I adopt this dynamic understanding of stereotypes to assess how incentives for politicians to behave in gender-stereotype-consistent ways may shift both over time and throughout the course of their careers. Third, gender role theorists suggest that individuals not only internalise the expectation that they should behave in gender-role consistent ways, but also internalise the expectation that others should do so too. I study both whether stereotypes inform how voters evaluate stereotype-(in)congruent behaviour by politicians, and how they inform how voters may differentially perceive the styles that men and women politicians express.

## **1.2 Substantive approach and contributions**

The papers in this thesis address whether stereotypes are influential in shaping the behaviour and attitudes of politicians and voters in the UK. In doing so, I contribute – both theoretically and empirically – to our understanding of how gender roles and stereotypes influence the behaviour of elites and voters.

In chapters 3 and 4, I study whether MPs conform to gender stereotypic behaviour in the UK House of Commons between May 1997 and March 2019. The House of Commons, the lower chamber of the UK parliament, is a classic majoritarian legislature with typically a single governing and one main opposition party (Lijphart, 2012). Over the period I study, there have been five general elections – 2001, 2005, 2010, 2015, and 2017 – and the period covers six parliamentary terms – 1997-2001, 2001-2005, 2005-2010, 2010-2015, and 2017-2019. Between 1997 and 2010, the Labour Party was the party

of government, between 2010 and 2015 the Conservative Party and the Liberal Democrats formed a coalition government, and between 2015 and 2019 the Conservatives formed a single party government, which, at the time of writing, remains the party of government. In chapters 3 and 4, the changing composition of parliament, and indeed career paths of women in parliament, is core to the theoretical arguments, which each rely in part on women's presence and status within the legislature. Throughout the period I study, there has been a concomitant rise in the number of women elected to the House of Commons. In figure 1.1, I show the number of women elected between 1992 and 2019 split by political party. As is clear from the figure, the 1997 General Election marked a landslide victory for the numeric increase of women, where 120 of the 650 MPs returned to the House of Commons (18.2%) were women, which saw the number of women double from the 60 women (9.2%) elected in the 1992 General Election. The number of women has since risen gradually over time and, by the end of the period of study, 208 women (32%) were elected to the House of Commons following the 2017 General Election.

While I am unaware of any over time survey questions about the changing endorsement of gender stereotypes in the UK specifically, recent evidence suggests that voters in the UK have become more gender-egalitarian in their attitudes (Taylor and Scott, 2018). Therefore, coupled with the increase in the number of women in politics over time in the UK, recent evidence suggests that gender roles have become notably less traditional over time, with particular changes occurring from 2008 onwards. In 1984, 37% of respondents to the British Social Attitudes survey disagreed that "a man's job is to earn the money; a woman's job is to look after the home and family", but this disagreement had risen to 72% in 2017 (Taylor and Scott, 2018, 1). Figure 1.2 documents changes in the levels of agreement and disagreement to this question between 1984 and 2017. Academic accounts also suggest that, in both the UK and further afield, support for traditional gender norms has eroded (Inglehart and Norris, 2003; Twenge, 1997a; Seguino, 2007). Further, between 1990 and 2010, voters in the UK, and Western Europe more broadly, have become significantly less likely to agree with the traditional division of social roles performed by men and women (Shorrocks, 2018).

Taken together, over the period I study, the presence of women in politics has increased markedly and the degree to which voters support traditional gender divisions is on the decline. While distinct in question and approach, the papers in this thesis each rest on the idea that stereotypes should be considered dynamic concepts that will shift as the presence and power of women in society and politics changes, as does public attitudes about these stereotypes. As such, the UK represents an interesting case to investigate the changing role of gender stereotypes in informing elite and voter behaviour.

In the first paper (chapter 3, co-authored with Jack Blumenau), I evaluate whether stereotypes

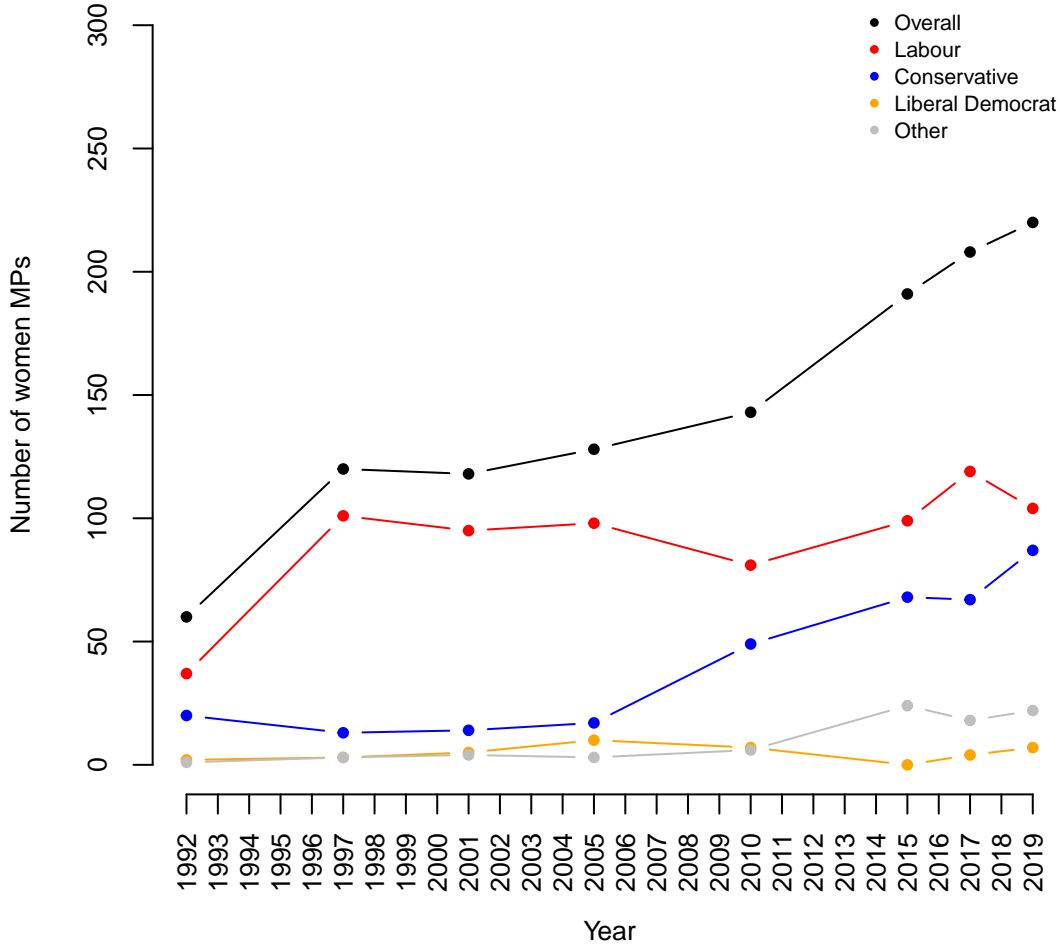


Figure 1.1: Number of women MPs by party, 1992-2019

serve as accurate behavioural descriptions of the ways in which UK MPs argue about politics in the UK House of Commons. Existing perspectives suggest that politicians face incentives to adopt styles which are consistent with gender stereotypes, and that these incentives induce stylistic differences between men and women’s speeches in parliamentary debate. Empirical evidence supports this perspective and shows that, compared to their men colleagues, women’s speeches are more emotional (Dietrich, Hayes and O’Brien, 2019; Gleason, 2020; Yu, 2013), less complex and jargonistic (Coates, 2015), less repetitive (Dahlerup, 1988), less aggressive (Kathlene, 1994), and use different types of evidence to support their arguments (Childs, 2004a). In the UK, existing work has shown that men and women’s styles differ:

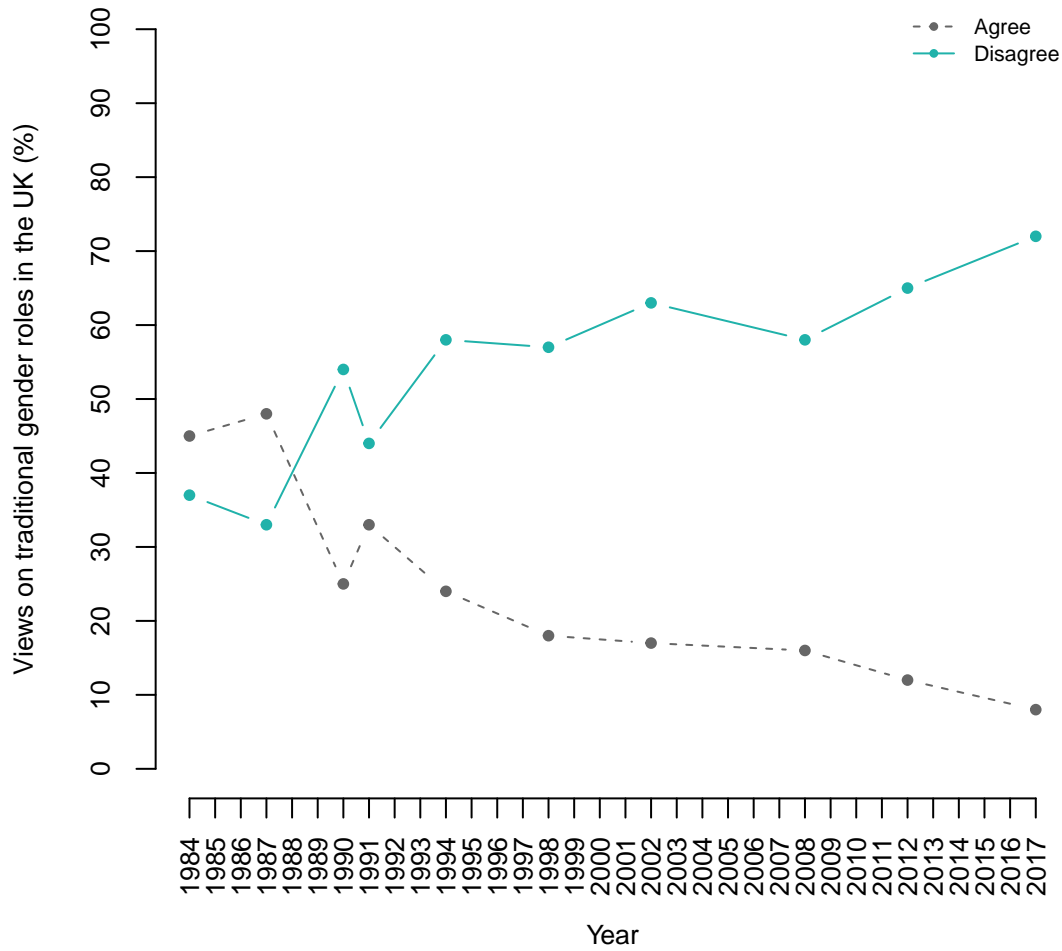


Figure 1.2: Percentage of respondents who agree or disagree with the statement “a man’s job is to earn the money; a woman’s job is to look after the home and family”, 1984-2017

women evidence arguments with personal experience, discuss policies in a concrete way, and are less adversarial than men (Hargrave and Langengen, 2021; Shaw, 2020).

However, existing work does not consider how gendered patterns of style have changed over time. Building on work that considers stereotypes to be *dynamic* constructs (Diekman and Eagly, 2000; Eagly and Wood, 2012; Eagly and Karau, 2002), I argue that over the past 25 years in the UK, several factors are likely to have decreased the degree to which UK MPs – and especially women – will conform to stereotype-consistent behaviours. First, politicians are drawn from a broader population which has it-

self diverged from stereotypical styles in recent years. As societal gender roles have changed, women in the public have been shown to demonstrate increasingly agentic behaviours across a wide set of contexts and countries (Leaper and Ayres, 2007; Twenge, 2001). Second, changes to the social roles played by women in public life, and in politics, have reduced the validity of gender stereotypes in the eyes of the public (as exemplified in figure 1.2). While attributes associated with men have remained relatively stable, masculine, agentic characteristics are increasingly ascribed to women (Eagly et al., 2020; Sendén et al., 2019). Consequently, I argue that voters are less likely to sanction women legislators for gender-incongruent behaviours now than in the past. Finally, the increased prominence of women in parliament (as documented in figure 1.1) and leadership roles (Blumenau, 2021c) is also likely to reduce the degree to which women politicians internalise expectations that they need to behave in “feminised” ways. Together, these arguments lead to a central behavioural prediction which I test empirically: that UK MPs will conform less to gender stereotypes now than in the past, and that women, in particular, will adopt styles that are further from feminine stereotypes over time.

To test this argument, I use novel quantitative text analysis approaches for measuring a diverse set of styles at scale in political speeches between 1997 and 2019. Analysing debates in the House of Commons, I show that in the early parts of my study period, politicians behave in a way that is broadly consistent with traditional gender stereotypes, where women primarily communicate in a manner that is consistent with “communal” stereotypes and men in ways that are consistent with “agentic” stereotypes. However, I show that these differences have reduced dramatically over the study period. In particular, I document a sharp convergence between men and women’s styles. Further, I show that the variation in gendered style usage has primarily resulted from women’s decreasing use of communal styles, and increasing use of agentic styles, over time. The implication of these findings is that traditional gender stereotypes now less accurately describe the reality of politicians’ behaviour than they once did. While previous work has largely addressed these questions at a fixed point in time, I provide an important contribution by documenting empirically how behaviour has shifted over time. I show that the descriptive validity of prominent stereotypes is considerably lower in the contemporary House of Commons than they were in the past.

Just as stereotypes are said to shape the styles that politicians express, so too are they said to shape the policy areas that men and women are expected to focus on. Women are said to pay greater attention to social policy issues that are related to their traditional roles as caregivers (Norris, 1996), and the long-standing stereotype that women are expected to be communal, caring, and compassionate (Eagly and Karau, 2002). As such, women are said to orientate towards policies that focus on caring for

others (Beutel and Marinin, 1995), and “compassion” issues such as poverty, education, children, family, and health policy (Shapiro, 1986). Empirical accounts suggest that in campaigns (Ennser-Jedenastik, 2017; Herrnson, Lay and Stokes, 2003) and in legislatures (Bäck, Debus and Müller, 2014; Bäck and Debus, 2019; Clayton, Josefsson and Wang, 2017; Schwindt-Bayer, 2006; Swers, 2002) women tend to pay greater attention to social policy issues. In the UK, work by Catalano (2009) and Bird (2005) suggests that women participate more in debates on health-related topics and question the government more on issues of women and gender. Existing perspectives have shown that, although context may mitigate the link (Childs and Krook, 2009; Hawkesworth, 2003), the issues that men and women prioritise are broadly in line with stereotypical expectations.

However, existing accounts have largely neglected to assess how this difference may shift over the course of politicians’ careers. As figure 1.1 shows, women have historically been underrepresented numerically in legislatures in both the UK and further afield (IPU, 2022), and have, on average, less experience in office than men (Lazarus, Steigerwalt and Clark, 2022). It has been challenging for studies to identify the implications that increased experience may have for both the behaviour of women legislators in office and for the representation of voters’ interests in the policy-making process. In the second paper (chapter 4), I build on literature on both the dynamics of legislative careers (Bailer and Ohmura, 2018; Bailer et al., 2022; Benedetto and Hix, 2007; Claessen, Bailer and Turner-Zwinkels, 2021) and political experience (Eggers and Spirling, 2017; Fulton, 2012), to understand the effect of politicians’ increased experience on their representational behaviour. Specifically, I argue that gender differences in the extent to which women politicians raise issues traditionally associated with women will be most pronounced among junior politicians, however, this difference will decrease with increased parliamentary prestige for several reasons.

First, at the beginning of their careers, politicians will gravitate towards the issue areas where they have the most prior experience. This issue selection will be shaped by a variety of factors but, for women, may also be shaped by their shared identity with women (Phillips, 1995; Wängnerud, 2009). Second, in the early stages, politicians need to gain credibility and competence in the eyes of voters and their fellow politicians and are therefore incentivised to prioritise issues where they will be ascribed the most authority and competence by others. As women have traditionally been perceived as better suited to, and to have greater expertise in, stereotypically “feminine” issue areas (Herrnson, Lay and Stokes, 2003; Sanbonmatsu, 2002), they will place greater emphasis on these issue areas. However, over time, politicians learn to navigate their roles and establish themselves and build their confidence in a wider variety of issue areas (Miquel and Snyder, 2006). Therefore, they will diversify the selection



of issues as they gain both greater experience in a wider set and more credibility for this work from voters and fellow politicians.

To evaluate this argument, I use quantitative text analysis approaches to detect the set of issues that politicians raise in House of Commons speeches. I find that, among junior politicians, women talk significantly more about “feminine” issues, however this gender gap decreases markedly as seniority increases. Further, I find that women politicians also raise women’s experiences to a greater extent than men, that they do so in debates on a wide range of issue areas, and that this does not change with increased political experience. Further, I find that some of the largest gender differences are concentrated in the more traditionally “masculine” policy areas, such as crime and policing or defence. The implications of these findings are positive, as they suggest that women diversify the sets of policy areas they focus on the longer they are in parliament, and that women continue to amplify women’s experiences throughout the course of their careers in debates across the policy spectrum.

The results I present in the first two papers suggest that the accuracy of stereotypes for describing politicians’ behaviour is contingent on both time and seniority. I show that, at the end of my study period, men’s and women’s communication styles have converged considerably. While politicians may now behave in a way that is less consistent with stereotypes than they once did, it remains possible that voters may both still perceive differences in elite behaviour and evaluate politicians differently based on their gender. In the third paper (chapter 5), I address this question directly by identifying the extent to which gender stereotypes shape voters’ attitudes towards politicians’ behaviour. When communicating with the public, politicians make regular use of a wide range of argumentation strategies such as empathetic appeals, emotional arguments, and aggressive attacks. While politicians expressing empathy, emotion, and aggression is unremarkable, certain politicians may face harsher penalties than others when they express, or do not express, these styles. I consider whether voters are biased in their perceptions and evaluations of the ways in which politicians communicate, and, consequently, whether voters’ evaluations of elite political communication are gendered.

While there is a rich literature on the influence of gender stereotypes on voter *evaluations* of politicians, this paper is the first to address whether voters differentially *perceive* politicians’ behaviour based on gender alone. I focus on styles for which the gender stereotypes literature has outlined clear expectations about men and women’s behaviour (Eagly and Wood, 2012; Schneider and Bos, 2019) and many of the styles assessed in chapter 3. Work to date examining how voters evaluate stereotype-(in)congruent behaviour has focused on isolated behaviours such as crying (Brooks, 2011), negative attacks (Cassese and Holman, 2018; Krupnikov and Bauer, 2014), willingness to compromise (Bauer,

Harbridge Yong and Krupnikov, 2017), or ambition (Saha and Weeks, 2022). By contrast, I make progress by focusing on how voters evaluate politicians' use of a wider range of styles that are consistent with both feminine "communal" stereotypes and masculine "agentic" stereotypes (Eagly and Karau, 2002). In doing so, I assess whether there are asymmetric standards in the degree to which men *and* women politicians are punished for stereotype-incongruent behaviours.

In a survey experiment, I present UK voters with speeches where the argument style and the gender of the politician delivering the argument are varied. Through these manipulations, I assess, first, whether voters perceive men and women politicians' behaviour differently based on stereotypical expectations, second, whether politicians incur backlash when violating gender-based norms, and third, whether differential perception of behaviour may itself explain this backlash. I find that style usage has important consequences for how voters evaluate politicians. Politicians are liked more when they are emotional or draw on anecdotes, however, they are simultaneously regarded as less competent when they use these styles. Contrary to the expectations from the gender stereotypes literature, I find no evidence that these evaluations are gendered: the effects of style on voters' evaluations of politicians' competence and likeability are equal for men and women politicians.

A commonly held perception is that women are punished when they adopt behavioural styles consistent with masculine agentic stereotypes (Fox and Oxley, 2003; Jalalzai, 2016). At the core of the idea that women politicians face double standards when they violate stereotypically expected behaviours is that voters actually hold these expectations for women's behaviour in the first place. However, as argued above, studies show that the public's perception of the validity of these stereotypes has shifted over time. Women have been seen as increasingly agentic over time (Eagly et al., 2020; Sendén et al., 2019). Voters in the UK have also become more gender-egalitarian in their attitudes over time (Taylor and Scott, 2018), and have become markedly less likely to support traditional gendered divisions in social roles (Shorrocks, 2018). Further, in chapter 3, I show that politicians in the UK have come to behave in a way that is less consistent with traditional gender stereotypes. Women politicians have decreasingly made use of "communal" styles and increasingly made use of "agentic" styles over time. The pessimistic assumption is that this behaviour change by women might be met by backlash from voters. However, the results I present in chapter 5 suggest that this may not materialise, as UK voters do not seem to unjustly penalise women politicians for stereotype-incongruent behaviour. Of course, without a study from 20 years ago to compare these findings to, it is not possible to know whether UK voters in previous eras did apply these descriptive stereotypes or punish women politicians for stereotype-incongruent behaviour. Yet, if voters no longer hold the same stereotypical expectations

about men's and women's behaviour, and politicians decreasingly behave in accordance with traditional stereotypes, it may not be that surprising to uncover that women politicians are not punished for behaviour that violates traditional stereotypes.

In sum, the first two papers provide new insights into the *dynamic* influence that stereotypes wield over politicians' behaviour, by assessing how these factors change over time and throughout the course of politicians' careers. The third paper focuses on voter behaviour and shows that while the use of styles consistent with stereotypes does inform how voters evaluate politicians, this does not seem to be gendered. Taken together, the papers in this thesis make several substantive contributions to our understanding of what shapes the behavioural strategies that political elites deploy, how voters engage with this behaviour, and to the changing role that gender stereotypes play in this process.

### **1.3 Methodological approach and contributions**

Beyond the substantive contributions described above, the papers also make use of innovative quantitative designs to understand what shapes the communication strategies that political elites deploy in the UK, and how voters perceive and evaluate this behaviour. I draw on two distinct methodological approaches to address these questions: quantitative text analysis and a survey experiment.

Chapters 3 and 4 focus on measuring the behaviour of political elites. To do so, I study the speeches that politicians deliver in the House of Commons to glean insights into the issues that politicians prioritise when participating in parliamentary debate, and the style in which they deliver these arguments. In both chapters, I focus on MPs' behaviour in one key aspect of their roles: their participation and behaviour in legislative debates. The behaviour of individual MPs is difficult to observe, and therefore measure, in many aspects of politicians' roles. Further, while individual-level behaviour is observable in parliamentary voting, the votes MPs cast are heavily influenced and controlled by their political party, and therefore voting behaviour largely does not provide meaningful information about politicians' behaviour. Instead, legislative debates represent a rare and useful source of panel data at the individual-level, and therefore enable researchers to study how the behaviour of MPs may change over time, or across debates on different topics or types, factors which are all crucial to my questions of interest. Plenary debates are a fundamental part of the legislative and representational process, where politicians can engage in arguments over policy, articulate their views, represent their constituents, appeal to new voters, and hold the government to account in a public forum. They are one of the most visible elements of a politician's job, receive coverage in the media, and are noticed by voters ([Camp-](#)

bell and Lovenduski, 2015; Proksch and Slapin, 2015; Rush and Giddings, 2011). They operate under relatively permissive rules and MPs are afforded a large degree of individual autonomy regarding the debates they participate in and, crucially, the content of their speeches (Bäck, Debus and Fernandes, 2021). The speeches that MPs deliver in the UK House of Commons provide meaningful insights into the styles and issues that individuals express and are an excellent source of data to assess my central questions of interest.

To measure both the styles politicians adopt in debate, and the substantive issues that they raise, I create behavioural measures that aim to take context and validity seriously. A common approach to measuring concepts, such as styles or topics, in text data is to assign political speeches a score on the basis of a predefined dictionary that aims to capture the concept of interest. However, dictionary-based approaches are highly domain-specific, as the words used to capture a concept in one context – say, parliamentary speeches – are likely to be different to those used to express the same concept in another context. As Grimmer and Stewart (2013, 268) argue, “applying dictionaries outside the domain for which they were developed can lead to serious errors”. For instance, one style that I assess in chapter 3 is whether politicians are aggressive in their speeches. When considering words that may be representative of aggression in, for instance, communication on social media platforms, we are likely to conjure up a set of words that is very different from how aggression is understood in the parliamentary context, where there are extensive rules on appropriate conduct and “unparliamentary” language. In essence, how we understand the concept of a particular communication style is likely to vary significantly depending on the context of interest.

I take this well-known critique of dictionary approaches seriously and introduce an alternative strategy that accounts for this critique. I develop measures that combine dictionaries with a locally-trained word-embedding model to construct domain-specific dictionaries that are able to capture my quantities of interest as they manifest *in the specific context of UK parliamentary debate*. Word-embedding models, which are now a mainstay of computer science research and are of increasing use in political science (Osnabrügge, Hobolt and Rodon, 2021; Rice and Zorn, 2021; Rodriguez and Spirling, 2022; Widmann and Wich, 2022) in essence rely on the idea that words which are used in similar contexts will have similar meanings. The core idea suggests that the meaning of a word can be intuited by examining the other words that appear close to it in text data. Or, as the commonly repeated maxim puts it, you can “know a word by the company it keeps” (Firth, 1957, 11). This approach, therefore, provides me with a measure of each of my styles and issues that is firmly grounded in the ways in which politicians debate in the UK House of Commons. The word-embedding approach allows me to detect

the style in which politicians communicate, and the substantive policy areas that they champion. In doing so, I can carry out a detailed exploration of gendered dynamics in UK House of Commons debates and emphasise the influence the stereotypes may wield over politicians' behaviour. These features of politicians' behaviour in office have become central quantities of interest to scholars in studying gendered and non-gendered behaviour alike. As such, this measurement approach will be of direct value to researchers interested in the study of politicians' speechmaking behaviour in legislatures globally.

Chapter 5 turns to measuring the behaviour of voters. Gendered perceptions and prejudice in politics can be troublesome to grapple with as it can be challenging to isolate gender as the factor that influences voters' judgements. In observational settings, it is difficult to ascertain which arguments voters receive and how they may evaluate these arguments. Experimental approaches, which have become increasingly common in the field of gender-based stereotyping (Campbell and Cowley, 2014; Campbell et al., 2019a,b; Saha and Weeks, 2022), offer a solution to this problem by enabling researchers to isolate the causal effect of gender on voter evaluations and perceptions of candidates or politicians (Mutz, 2011). This is of particular importance for my central questions, as I am interested in identifying whether voters differentially perceive politicians' style usage based on gender alone, and whether voters differentially evaluate men and women politicians depending on the styles they express.

To test whether stereotypes inform voters' judgements, I designed a vignette survey experiment, which was fielded by YouGov to approximately 1,700 members of their UK online panel. In the experiment, respondents were tasked with reading a short argument delivered by a fictitious MP. Respondents were assigned to different treatment conditions where various attributes were randomised. In particular, the styles, policy areas, and gender of the MP. To ensure the speeches were similar in spirit to the kinds of arguments MPs deliver in the UK parliamentary context, the texts of the speeches were informed by searching for debates on similar policy areas recorded in Hansard. When writing the speeches, the basic structure remained the same within a policy area, however, I added words and sentences that were deemed highly representative of the style types in the parliamentary context by selecting the highest-weighted words and sentences according to the word-embedding based approach described above. Overall, my aim was to ensure that the speeches appropriately mimicked UK House of Commons language and that the styles were representative of the same stylistic concepts studied in chapter 3.

Considered together, the papers in this thesis use creative and theoretically informed innovative quantitative designs to assess how gender stereotypes inform the way politicians behave, and how voters engage with this behaviour.

## 1.4 Implications for gender and politics

Stereotypes are said to inform both the way that politicians behave in office, but also shape voters' judgements about this behaviour. Together, my results suggest that women in parliament do not need to conform to behavioural stereotypes, and voters do not hold them to stereotypical standards. The substantive conclusions from this research therefore provide a correction to, and potential updating of, our collective understanding of how men and women may behave in office, and how voters engage with this behaviour. Ensuring that our assumptions about sexist attitudes and stereotypes in politics are grounded in empirical reality has important implications for politicians' experiences in office for several reasons.

First, when running for office, women are now no less likely to win than men (Lawless and Fox, 2010), indeed, women may now even experience a slight electoral advantage (Schwarz and Coppock, 2022). Evidence suggests that “when women run, women win” (Carroll and Sanbonmatsu, 2013). Gender gaps do, however, remain in women's likelihood to run for office (Bernhard, Shames and Teele, 2021; Lawless and Fox, 2005), the length of their political careers (Lazarus, Steigerwalt and Clark, 2022), and their likelihood to be promoted to higher office (Yildirim, Kocapinar and Ecevit, 2021). While the decision to run for office is likely driven by a plethora of factors, which I reflect further on below, women's reluctance to enter politics may, in part, be due to the assumption that voters will hold them to different standards. The testimonies of incumbent women suggest that they express concern about the influence of stereotypes and bias (Dittmar, Sanbonmatsu and Carroll, 2018), and report feeling more pressure and anxiety in their roles than their men colleagues (Erikson and Josefsson, 2019; Erikson and Verge, 2022). Women are therefore aware that stereotypes exist and believe that they may be upheld by voters. Expectations about incurring penalties for stereotype-incongruent behaviour may restrict the types of individuals who run for office and introduce damaging restrictions on the ways in which politicians perceive it is acceptable for them to behave. Instead, I show both that politicians are now less likely to conform to stereotype-expected behaviours than they were in the past, but also that voters may not always punish women any more than men for stereotype-incongruent behaviour. Highlighting that voters may not be as biased towards women as anticipated is important as it may help to encourage more women to run for political office.

Second, a source of bias that may restrict women's accession to higher office is that party elites may not put women forward for fear that *others* will harbour negative attitudes towards them. Bateson (2020) identifies the concept of “strategic discrimination” and argues that political leaders or voters

may minimise the abilities or likelihood of success of particular candidates – notably, women or ethnic minorities – because they anticipate that others will likely not vote for or support them. Because they want their candidates and party to be successful, they will not put women forward for fear that they will not win. In essence, bias against women in politics can occur when individuals or parties *overestimate* the degree to which others are biased. This may be of particular concern for women’s promotion to leadership positions. It is commonly assumed that stereotypically “masculine” styles of leadership are preferred for senior political positions (Fox and Oxley, 2003; Jalalzai, 2016), and that when women try and adopt these styles, they risk violating feminine stereotypes, and may lose out. However, I show that the reality for women politicians may be more positive than this, as the expectation that they must avoid these behaviours for fear of backlash is somewhat misguided. Therefore, it is important to document that voters may not be as biased towards women as anticipated, as it may help to further reduce bias and ensure that women are not held back from leadership positions.

Finally, as the examples of women’s experiences from the opening paragraph demonstrates, sexism persists as a serious issue in politics. The findings I present are not intended to suggest that the influence of stereotypes has entirely disappeared, nor that sexism is no longer at play. Rather, I ground two key areas in which stereotypes are said to be influential in empirical reality – on the manner in which politicians behave in parliament and how voters engage with this behaviour – and show that, in these contexts, the power of stereotypes may now have declined. Women politicians may, however, continue to be sanctioned for their behaviour by other sources. In particular, as recent press coverage by the Mail on Sunday accusing Labour MP Angela Rayner of “distracting” senior politicians with her legs has shown, the media may play an important role in the perpetuation of sexism in politics. Men have been shown to receive more coverage overall than women (Banwart, Bystrom and Robertson, 2003; Bystrom, Robertson and Banwart, 2001; Falk, 2008), yet, when women are covered, this coverage tends to be both more negative (Heldman, Carroll and Olson, 2005; Kahn, 1992; Ward, 2016), more focused on traits as opposed to issues (Aday and Devitt, 2001; Kahn, 1994; Ross et al., 2013), and to understate women’s abilities relative to men’s (Bauer and Taylor, 2022). Therefore, if the media does indeed frame women’s behaviour in a more negative light than men’s, this may in turn feed into how voters evaluate politicians even if voters’ direct judgements are not themselves gendered.

Of course, it is vital to recognise that gender stereotyping is only one potential form of bias that may make women’s experiences in political life more challenging than men’s experiences. For instance, surveys conducted with UK MPs suggest that the role of an MP may be one that is incompatible with parenthood. Examining the gaps between men and women, Campbell and Childs (2014) documented

that while only 28% of men MPs reported having no children, the equivalent figure for women was 45%. Therefore, at the time the survey was conducted, only just over a half of women in parliament were both MPs and parents. Comparing these figures to Office for National Statistics (ONS) data, they argue that the percentage of women MPs without children is more than twice the average among women in the public (Campbell and Childs, 2014, 489). There is a notable disparity between the percentage of working mothers in parliament compared to working mothers in fields beyond politics. Although recent evidence suggests that gap may be declining over time, with 39% of women MPs compared to 30% of men MPs reporting having no children in 2017.<sup>2</sup> Researchers working on gender sensitive parliaments have often raised the lack of family-friendly working practices in the UK Parliament as a significant hurdle for women politicians to serve their roles while also taking on motherhood (Childs, 2016). While important recent reforms have helped increase the practicality of balancing parenthood and a position in parliament, such as the introduction of proxy voting for MPs on baby leave,<sup>3</sup> factors such as reforming parliamentary sitting hours and ensuring formal codes of behavioural conduct are effectively enforced are still needed.<sup>4</sup> As such, the lack of family-friendly working practices in parliament represent a barrier to balancing the job of an MP with caring responsibilities.

The rise in the harassment and intimidation of candidates and politicians both online and offline also presents a very real challenge for increasing women's entry into politics. The brutal murder of Labour MP Jo Cox in 2016, while she was campaigning for the UK to remain in the European Union, prompted serious concerns about the personal safety of politicians. Ahead of the 2019 General Election, several women politicians, including Nicky Morgan, Heidi Allen, and Amber Rudd, also decided not to run for re-election, declaring abuse, harassment, and intimidation as some of their key reasons for doing so. More broadly, surveys conducted with candidates in recent UK elections suggests that women are far more likely to report experiencing abuse and concern for their personal safety than men (Collignon and Rüdiger, 2020), and that women also disproportionately experience sexual abuse both online and offline (Collignon and Rüdiger, 2021). Work by Collignon, Campbell and Rüdiger (2022, 35) suggests that the harassment and intimidation of parliamentary candidates may even be worsening over time, with 45% of women in 2017 compared to 58% in 2019, and 35% of men in 2017 compared to 44% in 2019, indicating that they had experienced some form of harassment while campaigning. With more than half of women reporting that they campaign in fear for their safety (Collignon, Campbell and Rüdiger, 2022, 35), the increasing experience of harassment, abuse, and intimidation in politics represents

<sup>2</sup>See, for example, "Where's Mum? She's in the House – Parents in Parliament", PSA, 29th March 2019.

<sup>3</sup>See, for example, "Victory! How I fought to get maternity leave for female MPs", *The Telegraph*, 25th September 2020.

<sup>4</sup>See, for example, "The Good Parliament: what kind of Speaker do we need?", *The Constitution Unit*, 1st October 2019.



a serious concern that particular types of individuals may be deterred from becoming, or continuing to be, active in political life.

Sexist attitudes and behaviour remain problematic in politics. The current working practices of the UK parliamentary system make maintaining the job of an MP, while balancing caring responsibility specifically, and a healthy private life more broadly, a serious challenge for certain types of politicians. Candidates and politicians, and in particular women, face growing concerns for their safety when carrying out key components of their political roles. In short, improving and equalising women's experiences in contemporary political life is an ongoing challenge that needs to be confronted. The work I present in this thesis provides an important contribution by shedding light on the areas where gendered biases and barriers in politics may have declined in recent years. However, these findings should be considered within this much wider context of women's experiences in politics. The results I present in the remainder of this thesis should motivate researchers in and out of academia to turn their attention to these other features of political life where gendered imbalances in politics remain rife.

## Chapter 2

# Literature Review

In this thesis, I ask whether UK politicians behave in accordance with traditional gender stereotypes, how the pressure to do so shifts by context and over time, and whether voters in the UK uphold these stereotypes when they perceive and evaluate politicians' behaviour. In chapters 3 and 4, I focus on how stereotypes inform the behaviour of elite politicians, in chapter 5, I turn to the behaviour of voters and assess whether stereotypes inform how they perceive and evaluate elite behaviour. In each of these chapters, I delve further into the literatures that underpin the specific questions and arguments. Here, I review the theoretical literature on how gender roles and stereotypes influence this behaviour, as well as the existing empirical work examining how these stereotypes may inform both the manner in which politicians carry out their roles, but also how voters engage with their behaviour.

I make several contributions – both theoretically and empirically – to the literature described below on the influence of gender roles and stereotypes on the behaviour of elites and voters. The first paper (chapter 3) evaluates whether stereotypes serve as accurate behavioural descriptions of the ways in which MPs debate in the UK House of Commons. While previous accounts have presented evidence broadly consistent with this perspective, I contribute to this work by considering, and testing empirically, how gendered patterns of behaviour may vary over time. The second paper (chapter 4) focuses on the degree to which stereotypes may guide the sets of issues that politicians raise and prioritise during their time in office. Past work has emphasised that women politicians may focus more on traditionally “feminine” policy areas and have presented findings broadly consistent with this. I contribute to this work by arguing that differences in the extent to which politicians raise issues traditionally associated with women will be largest among junior politicians who lack experience in the eyes of the public and their colleagues, however this will decline with increased political experience. In the third paper (chapter 5), I consider the influence stereotypes are said to wield over how voters engage with politicians' behaviour. While there is a robust literature on the influence of gender stereotypes

on voter evaluations in politics more broadly, I contribute to this work by assessing whether voters differentially *perceive* the behaviour of men and women politicians. I consider, first, whether voters' perceptions of the styles themselves are gendered, second, whether politicians incur backlash when violating gender-based norms, and third, whether differential perceptions of the styles themselves explain this backlash.

## 2.1 The formation of gender stereotypes

Stereotypes are said to emerge from divisions in traditional societal gender roles. Underpinning the work on gender stereotypes is the central idea that society classifies individuals into men and women based on their biological sex, and these categorisations are associated with strong expectations of how people of different genders can and should behave (Eagly and Karau, 2002). These societal divisions emerge from the division of labour between men and women, which originates from both cultural and socioeconomic factors, such as women's role as homemakers and caregivers, and men's role as the financial breadwinners and providers for the family, but also due to physical differences, where women are childbearers and men exert physical prowess (Eagly, 2013; Eagly and Wood, 2012). These divisions come together to form a societal norm for how men and women are expected to behave. From a young age, men and women therefore encounter various experiences through education, workplaces, and social interactions that reinforce and reflect these behavioural differences in many small but cumulative ways. Eventually repeated exposure to these experiences can lead men and women to *internalise* these expectations. Women learn to behave in a way that is consistent with the "feminine" gender role that women should be accommodating, consensual, and emotionally sensitive, and men to behave in a way which is consistent with the "masculine" gender role that men are tough, assertive, and dominant (Schneider and Bos, 2014, 2019).

These stereotypical expectations inform the manner in which people behave, such that women are more "affectionate, helpful, kind, sympathetic, interpersonally sensitive, nurturant, and gentle" and men are more "aggressive, ambitious, dominant, forceful, independent, self-sufficient, self-confident, and prone to act like a leader" (Eagly and Karau, 2002, 574). However, they can also inform the attitudes and priorities of men and women, which are related to the personality traits, such as agreeableness, empathy, and assertiveness, that men and women are said to have (Eagly, 2013; Huddy, Cassese and Lizotte, 2008). Women are said to orientate towards policies that focus on caring for others (Beutel and Marinin, 1995), and "compassion" issues such as poverty, education, children, and health policy

issues (Shapiro, 1986).

The stereotyping of men and women based on their gender-roles is common practice in both the social (see Fiske and Stevens, 1993, for an overview) and political realm (e.g., Barnes and Beaulieu, 2019; Bauer, 2020a; Karpowitz and Mendelberg, 2014; Koch, 2000, and discussed in more detail below). A key distinction is drawn in this literature between *descriptive* stereotypes and *prescriptive* stereotypes (Heilman, 2001). Descriptive stereotypes relate to describing how men and women *typically* behave, for example, the idea that “women like to take care of others” or that “men are strong leaders” (Gill, 2004). Distinct from descriptive stereotypes, prescriptive stereotypes instead outline how men and women are *expected* to behave, such that “women like to care for others” becomes “women *should* care for others” and “men are strong leaders” becomes “men *should* be strong leaders” (Gill, 2004, 620). These prescriptive stereotypes outline the behavioural norms by which men and women are therefore expected to conform to avoid incurring criticisms, penalties, or backlash from others.

Descriptive and prescriptive stereotypes both operate in slightly distinct ways. As descriptive stereotypes provide beliefs about how men and women typically behave, this can lead to assumptions being made about the characteristics of individuals based on their gender. For example, the assumption that a woman enjoys caring for others simply because of her gender is a typical example of a descriptive stereotyping effect. Prescriptive stereotypes, however, help to uphold social norms about how men and women are expected to behave. If men and women then behave in a way that is congruent with their gender roles then this behaviour is deemed acceptable, however, if they instead behave in a way that is incongruent with their gender roles then this can result in negative backlash. For instance, a woman may experience backlash from others if she violates the prescription that she *should* like to care for others. Instead, because this prescription does not apply to men, they will not suffer for any behaviour which is incongruent because they are not violating a stereotype (see Cialdini and Goldstein, 2004; Gill, 2004, for a review of this literature). In short, descriptive stereotyping refers to the process by which we associate particular behaviours with men and women (e.g., men and assertiveness or women and emotionality), whereas prescriptive stereotyping instead relates to the idea that individuals are held to these stereotypes and may incur negative sanctions by others if they fail to do so.

According to gender role theory, how might men and women in society behave? A key distinction in the social psychology literature, originally introduced by Bakan (1966), is drawn between *communal* stereotypes, which are typically associated women, and *agentic* stereotypes, which are typically associated with men (Diekmann and Eagly, 2000; Rucker, Galinsky and Magee, 2018; Sczesny, Nater and Eagly,

2019). Broadly speaking, communality and agency describe how people relate to the social world, and represent a framework for understanding values, traits, motives, and behaviour. Communal stereotypes are primarily concerned with the welfare of and caring for other people and their well-being, for instance, the idea that women are “affectionate, helpful, kind, sympathetic, interpersonally sensitive, nurturant, and gentle” (Eagly and Karau, 2002, 574). Agentic stereotypes are instead orientated towards goal attainment, assertiveness, control, and confidence, for instance, the idea that men are “aggressive, ambitious, dominant, forceful, independent, self-sufficient, self-confident, and prone to act as a leader” (Eagly and Karau, 2002, 574). Across a range of contexts (Best and Williams, 1990), women are said to be warm (Broverman et al., 1972) or expressive (Spence, Helmreich and Stapp, 1974), whereas men have been described as competent and rational (Broverman et al., 1972; Eagly and Steffen, 1984) or instrumental (Spence, Helmreich and Stapp, 1974).

The literature discussed here suggests that gender roles become self-reinforcing, as societal beliefs about the typical behaviours of men and women lead to ever-more entrenched patterns of gendered behaviour. However, gender role theorists also suggest that stereotypes are not necessarily fixed but can instead be considered to be dynamic concepts (Diekman and Eagly, 2000; Eagly and Wood, 2012; Eagly et al., 2020). As described above, the stereotypical characteristics of a group are related to the activities associated with their typical roles in society. For instance, because women have traditionally occupied domestic roles in the home where they provide care to others, “caring” as a characteristic became stereotypic of women. However, while gender divisions in society still exist, the kinds of roles men and women hold have shifted over time, and the characteristics associated with men and women may have shifted too. As Diekman and Eagly (2000, 1172) argue, “to the extent that gender stereotypes reflect observations of men and women in social roles, perceivers should believe that their characteristics change as their distributions into social roles change”. While divisions still exist in domestic labour (Bianchi et al., 2014), and large gaps persist in key divisions of labour in the household (Weeks, 2022), gender divisions in the workforce have converged considerably over time (Ortiz-Ospina and Tzvetkova, 2017), as women have increasingly entered the labour force in a diverse variety of roles (Blau, Brummund and Liu, 2013; Lippa, Preston and Penner, 2014). As women have moved away from just the domestic sphere, which is associated with communality, and increasingly taken up roles in the workforce, the stereotypes associated with women have taken on characteristics associated with these environments (Eagly et al., 2004), which Eagly and Steffen (1984) show are more agentic and less communal.

Diekman and Eagly (2000) argue, and confirm in a series of experiments examining perceptions of

gendered roles and personalities over time, that individuals may therefore see that gender differences are declining over time because of increasing similarity in the societal positions that men and women hold. Further, they argue that the stereotypes associated with women should change in particular because the largest societal shifts have occurred in changes in women's behaviour. In a meta-analysis, [Twenge \(1997b\)](#) finds large gains in the degree to which women self-report association with agentic traits over time, and, while [Twenge \(2001\)](#) reports few shifts in men's behaviour, women are shown to increasingly relate to the agentic personality traits of assertiveness and dominance. These differences are reflected in speech-based differences, where the tendency for men to be more assertive has declined over time ([Leaper and Ayres, 2007](#)).

In a recent study, [Eagly et al. \(2020\)](#) conduct a meta-analysis of opinion polls on communal and agentic traits and their association with men and women between 1946 and 2018. They argue that as women still have greater occupation in communal, household, and family-making roles, then women will still be associated with communal stereotypes. However, as women have also increasingly moved into traditional masculine-dominated occupations such as lawyers, managers, and leadership positions more broadly ([Carli and Eagly, 2017](#)), they will increasingly be associated with agentic stereotypes too. They find that while the association between women and communality has remained relatively stable, women have been perceived as increasingly agentic over time. While these studies suggest that differences between men and women are decreasing for masculine, agentic attributes, there are considerably smaller shifts in feminine, communal attributes. While women have increasingly entered in traditionally masculine roles and occupations, men have not markedly increased their participation in traditionally feminine roles, and their positions in the domestic sphere are still primarily as the financial provider for the family ([Cejka and Eagly, 1999](#); [Queneau, 2006](#)). Women therefore still disproportionately specialise in communal roles, and meta-analytic studies suggest that self-reported communal traits have also remained relatively stable over time ([Twenge, 1997b](#)).

From the perspective of gender role theory, stereotypical ideas of the types of behaviour associated with men and women is ultimately rooted in gendered societal divisions. If men hold leadership positions and high-powered roles in the public sphere, they come to be associated with agentic characteristics, just as women's concentration in the care-taking roles in the private sphere leads to their association with communal characteristics ([Wilde and Diekman, 2005](#)). Therefore, as social roles in society shift, the characteristics that are ultimately associated with men and women may shift too. The implication of these arguments is that group stereotypes should be considered *dynamic*. Stereotypes will inform how men and women in society are expected to behave, however, the content of these

stereotypes may shift over time as the distribution of men and women into different social roles also shifts.

While much of the literature described here relates to the experiences of men and women in society more broadly, I turn next to reviewing how prior work has applied these theories to the *political* realm specifically.

## 2.2 The application of gender roles to politics

Women remain underrepresented in political institutions globally holding, on average, only one-quarter of legislative seats (IPU, 2022). Historically studies have found that women are less likely to be selected as candidates for political office (Lawless and Fox, 2005; Norris and Lovenduski, 1995), and, upon running are less likely to win elections than men (Darcy, Welch and Clark, 1994; Fox and Oxley, 2003; Lawless, 2004). Further, when running for office, women of equivalent qualifications are less likely to perceive themselves as sufficiently qualified (Lawless and Fox, 2012), and instead more likely to underrate their own competence (Bauer, 2020a) or shy away from competition (Buser, Niederle and Oosterbeek, 2014). Over time, these gender gaps have declined, and women's experiences have improved. There has been significant progress in equalising the resources, skills, and qualifications men and women have when running for office (O'Brien and Rickne, 2016; Profeta and Woodhouse, 2022; Weeks and Baldez, 2015). Recent studies suggest that women candidates may even be *more* qualified than men with respect to political experience, professional profiles, and education levels (Fulton, 2012). Further, studies find that there is less evidence of voters holding an outright preference for men over women (Campbell and Cowley, 2014; Campbell et al., 2019b; Dolan, 2014; Vivyan and Wagner, 2015), and that women now win at equivalent rates to men (Lawless and Fox, 2010; Schwarz and Coppock, 2022).

Yet, despite this growing equality in the experiences of women running for office and in electoral outcomes at the ballot box, it remains the case that women are underrepresented. Identifying why women have yet to reach parity in political office is a key question in the literature on gender and politics, and approximately two approaches have been taken to address this. The first strand of research examines the factors that influence women's decision to run for office. This has been addressed through questions such as whether socialisation through career paths (Welch, 1977), psychological elements such as political ambition (Fox and Lawless, 2004; Shames, 2017) and aversion to risk (Kanthak and Woon, 2015), or institutional barriers such as the type of electoral system (Iversen and Rosenbluth,

2010; Kittilson and Schwindt-Bayer, 2012) or candidate selection practices (Burrell, 1994; Norris and Lovenduski, 1995) affect women's likelihood to run for office.

The second strand of research focuses instead on what causes individuals to support women candidates. A vast body of work has studied how the power and prominence that gender stereotypes may wield over candidates' and politicians' behaviour once they decide to run for office (Schneider and Bos, 2014, 2019; Teele, Kalla and Rosenbluth, 2018). As politicians are selected from the wider population, they too will likely have internalised the expected behaviours associated with their gender roles prior to entry into office. Further, many studies have shown that women politicians in particular are subject to pressures to conform to stereotypes and may incur penalties from voters (Bauer, 2015a, 2020a; Boussalis et al., 2021; Cassese and Holman, 2018; Everitt et al., 2016; Mo, 2015), their colleagues (Bauer, Harbridge Yong and Krupnikov, 2017; Gleason, 2020; Gleason, Jones and McBean, 2019; Kathlene, 1994), and party hierarchy (Yildirim, Kocapinar and Ecevit, 2021) when they violate these stereotypes. The literature on gender stereotyping in politics suggests that the incentives men and women in the public may have to adhere to gender-role-consistent behaviour are also present for politicians.

However, in addition to pressures to comply with behaviours that are consistent with their gender, politicians are also expected to balance this with the behaviours consistent with *leadership* stereotypes. Non-gendered accounts of political leadership suggest that leaders are stereotypically expected to be assertive, competitive, dominant, persistent, driven, and confident (Epitropaki and Martin, 2004; Offermann, Kennedy Jr. and Wirtz, 1994), and work studying the behaviour of leaders suggests that this matches to politicians' behaviour, at least in some contexts (Elgie, 2015; Foley, 2013; Northouse, 2016; Weinberg, 2021b). However, as men have not only legally been able to serve in political office for longer than women but have also disproportionately held legislative seats and leadership positions (Teale, 2018), the stereotypes and behaviours associated with leadership have been shaped by the behaviour and priorities of men. In a meta-analytic study, Koenig et al. (2011) assess the extent to which stereotypes of leaders are culturally masculine. They find that stereotypes of leadership are masculine, and that leaders are perceived to be like men, and not like women, as well as more agentic than communal, and as more masculine than feminine. However, they find some evidence that this has shifted over time in the more recent set of studies they assessed (Koenig et al., 2011, 634-635).

Further, just as men's occupation of political roles may have shaped the expected behaviours of leaders, political arenas and systems have historically been dominated and shaped by men and therefore "embedded with masculine norms" (Schneider and Bos, 2019, 184). Feminist institutionalists have long argued that the gendered nature of institutions interacts with men and women's behaviour



(Erikson and Josefsson, 2019; Erikson and Verge, 2022; Krook and Mackay, 2011; Mackay, Kenny and Chappell, 2010; Mackay, 2008; Puwar, 2004; Ross, 2002). Political institutions are said to produce, maintain, and reproduce gendered experiences and are “far from neutral” in how they adapt to legislators of different genders (Childs and Webb, 2012, 32). Hawkesworth (2005, 147) argues “embedded in organizational rules, routines, and policies, gender power normalizes male dominance and renders women, along with their needs and interests, invisible”. As men are the traditional occupants of these institutions, institutional behavioural norms have been shaped in such a way that they are compatible with “masculine” behavioural norms, however, may be incompatible with “feminine” norms of behaviour. Franceschet (2011, 66) therefore argues “women may respond by disavowing distinctly feminine (and feminist) concerns, instead favouring the style and substantive issues of the dominant group.”

Therefore, in addition to the prescriptions from their gender roles, politicians also must balance the traits associated with both leadership *and* political institutions. Men are the traditional occupants of both political leadership positions and institutions. The stereotypes associated with masculine behavioural norms are congruent with the stereotypes associated with these roles too. However, for women politicians, the incongruence between traditional “feminine”, leadership, and institutional behavioural stereotypes may present a challenge to balance a complicated array of behavioural expectations. The implication, therefore, is that women may face difficulties in satisfying the behaviours associated with both their gender roles and the behaviours deemed appropriate and preferable for political leadership positions. For men, however, this poses little challenge as the traditional “masculine” traits of political leadership – such as assertiveness, competition, ambition, and outgoingness (Koenig et al., 2011) – are the same set of traits considered normatively desirable for men.

Political scientists have dedicated significant attention to assessing whether men and women politicians are incentivised to behave in a way that is consistent with the stereotypes associated with the gender. In chapter 3, I focus on this with respect to the *styles* that politicians express, and in chapter 4, I study the sets of *issues* that politicians dedicate their attention to.

### **2.3 Style-based behavioural differences**

On which dimensions of political behavioural styles should we expect gender differences? Style can be broadly defined, not just with respect to speech, but also in the manner in which politicians carry out their wider political roles (Fenno, 1978). The gender stereotypes literature suggests that individuals face incentives to conform to a complicated mix of behaviours. As introduced above, of central concern

in the social psychology literature is the distinction between *communal* characteristics, which are associated with women, and *agentic* characteristics which are associated with men (Eagly and Wood, 2012; Schneider and Bos, 2019). These labels are heuristics for clusters of behavioural attributes: communal characteristics are said to include being “affectionate, helpful, kind, sympathetic, interpersonally sensitive, nurturant, and gentle”, while agentic characteristics include being “aggressive, ambitious, dominant, forceful, independent, self-sufficient, self-confident, and prone to act like a leader” (Eagly and Karau, 2002, 574). In political science, these concepts have been operationalised using a diverse set of indicators. So far, however political scientists have tended not to map the empirically observable aspects of style upon which they focus with the theoretical concepts of communality and agency which emerge in the gender stereotypes literature. A central goal of chapter 3 is to capture a broader set of styles that might more fully reflect the *theoretical* ideas of communality and agency that emerge in the social psychology literature. To facilitate this aim, I provide some structure by bringing together the theoretical concepts of communality and agency with the various ways in which political scientists have operationalised gendered style usage.

I begin by considering the various indicators of style that fall under the rubric of *communal* characteristics. As a result of both socialisation (Bear and Woolley, 2011) and also as a strategic device to overcome structural gendered barriers and legislative marginalisation (Krook and O’Brien, 2012), women supposedly engage more in collaborative behaviour, which can capture the act of working with others to craft policy, to network, to organise meetings, to exchange ideas and gather information, and to encourage colleagues to participate and speak out in meetings, floor debates, or committees. Empirically, studies have assessed collaborative behaviour through interviews with politicians (Barnes, 2016; Childs, 2004b, 2013; Tolleson-Rinehart, 2001), bill co-sponsorship behaviour (Barnes, 2016; Holman and Mahoney, 2018; Holman, Mahoney and Hurler, 2021; Swift and Van der Molen, 2021), network analysis (Wojcik and Mullenax, 2017), and even participation in social activities in the US Congress such as congressional baseball games (Lawless, Theriault and Guthrie, 2018). Overall, these studies have presented evidence that women do engage more in collaborative behaviour than men.

One key way in which the political science literature has studied communality is with respect to women being more emotional, and, in particular, more positive and empathetic than men. Primarily, empirical work has investigated emotionality through an analysis of politicians’ speeches (Bligh and Kohles, 2008; Gleason, 2020; Jones, 2016; Yu, 2013) and shown that women’s speeches are characterised by a higher level of emotionality than men’s, and make greater use of positive emotional language such as expressing happiness or joy (Yu, 2013). Further, other studies have moved beyond an assessment

of legislative speeches, to instead investigate vocal pitch, tone, or facial expressions (Boussalis et al., 2021; Dietrich, Hayes and O'Brien, 2019).

Another stylistic indicator associated with communality is the idea that women draw on different sorts of evidence than men. Women's argumentation styles are said to "start from the personal" (Bochel and Briggs, 2000, 66), and make greater use of personal tone, experiences, and examples (Dahlerup, 1988), to consist of narrative, analogies, and anecdotes (Banwart and McKinney, 2005; Dahlerup, 1988), and emphasise "lived experience" (Blankenship and Robson, 1995, 359). The testimonies of politicians support the idea that women make greater use of anecdotes and personal experience (Childs, 2004b; Sones, Moran and Lovenduski, 2005), as do studies that measure gender differences through political speech data (Broughton and Palmieri, 1999; Hargrave and Langengen, 2021; Johnson, 2005). Therefore, for the features of style related to communality, such as collaboration, networking, encouraging others, emotionality, positivity, and human interest, the existing empirical evidence seems to support the expectations of the theoretical literature.

Political scientists have also investigated various indicators of style that fall under the rubric of *agentic* characteristics. While women are said to rely more on personal narrative, men are said to rely on "analytical, organised, and impersonal" language (Jamieson, 1995, 76), that relies heavily on empirical evidence, statistics, and figures (Robertson et al., 1999). Studies reporting on politicians' testimonies suggest that men politicians pay greater attention to "scientific research" and "back[ing] everything up with figures" (Childs, 2004b, 180), while other studies suggest the opposite dynamic: that women may use *more* facts (Barnes, 2016). Studies examining political speech data have shown no evidence that men make greater use of facts (Mattei, 1998), and that women may even make greater use of facts and statistics (Hargrave and Langengen, 2021).

Men are also said to use language that is more complex, formalistic, and jargonistic, while women are thought to be accessible and clear (Coates, 2015). Similarly, men are thought to have a tendency towards repetition (Gomard and Krogstad, 2001) and "always do want to say it again" (Childs, 2004a, 6). Studies on the complexity and repetitiveness of speech find men do have a tendency towards babble and jargon (Childs, 2004a; Gomard and Krogstad, 2001), whereas women are said to be "less long winded" and to "repeat themselves less" (Gomard and Krogstad, 2001, 159). Further, men are accused of not just repeating their own arguments again and again, but also "duplicating" from one another. Further, women are shown to use a "simpler and more concrete" language (Karvonen, Djupsund and Carlson, 1995, 346), that is more explicit and concise (Holmes and Meyerhoff, 2003; Lin and Osnabrügge, 2018; Thelander, 1986).

Of central focus is the idea that men are said to be more aggressive and engage in greater attacking language in their communication, whereas women are said to avoid a combative and aggressive style (Bochel and Briggs, 2000; Karpowitz and Mendelberg, 2014). Women MPs suggest that they preferred a “kinder, gentler politics”, that was characterised by co-operation rather than conflict, collaboration rather than hierarchy, and honesty rather than sleaze” (Norris, 1996, 93). Further, women claim they are “more willing to listen to the other side”, are “less adversarial [and] better team players” (Bochel and Briggs, 2000, 66-67), preferred a “less combative and aggressive” style, and do less “standing up and shouting on the floor of the House” (Childs, 2004a, 5). Women have been shown to interrupt less than men (Ban et al., 2022; Cameron and Shaw, 2016, 2020; Jacobi and Schweers, 2017; Karpowitz and Mendelberg, 2014; Kathlene, 1994, 2001; Mendelberg, Karpowitz and Oliphant, 2014), to engage less in political insults or point-scoring (Banwart and McKinney, 2005), and to make speeches that are overall less adversarial and aggressive (Grey, 2002; Kathlene, 1994; Hargrave and Langengen, 2021; Shaw, 2000, 2002, 2006, 2020). Further, while women are said to make greater use of emotionality overall, they are said to avoid negativity for fear of being penalised (Cassese and Holman, 2018), whereas men are said to make greater use of emotional negativity such as expressing criticism or disappointment (Brooks, 2011). However, empirical investigations suggest that women may make greater use of negative emotion in campaigns and advertisements (Kahn, 1992; Robertson et al., 1999).

Relatedly, men are also said to be more authoritative and dominant than women. Political science studies have typically operationalised this through assessing whether women are less likely to participate than men. Overall, men have been shown to participate more in parliamentary debates (Brescoll, 2011; Pearson and Dancy, 2011a; Wang, 2014), although these differences have been shown to become more equitable depending on the composition of the group (Ban et al., 2022; Karpowitz and Mendelberg, 2014; Karpowitz, Mendelberg and Mattioli, 2015a,b), presence of a woman leader (Blumenau, 2021c), or the type and topic of debate (Bäck, Debus and Müller, 2014; Bäck and Debus, 2019; Catalano, 2009; Fernandes, Lopes da Fonseca and Won, 2021). Therefore, for some styles consistent with agency, such as complexity, repetition, interruptions, aggression, and participation differences, the empirical findings align with the theoretical expectations. However, for other features of style, such as the use of facts and statistics or negative emotion, the empirical findings reveal no evidence of differences in the expected direction.

Throughout the literature on gendered style usage in politics, there are some common limitations which I seek to make progress on. First, most studies measuring political style by observing legislator behaviours are reliant on very small samples of speeches or debates (e.g., Broughton and Palmieri, 1999;

Kathlene, 1994). In chapter 3, I provide evidence from many thousands of debates and speeches over a large time period, and thus significantly expand the scope of my analysis relative to other studies. Second, although the studies that have taken quantitative approaches have expanded the population of speeches (e.g., Bligh and Kohles, 2008; Jones, 2016; Yu, 2013), these studies examine only a narrow range of stylistic indicators. As my review of the literature reveals, there are many dimensions on which we might expect gendered differences to exist. Quantitative studies that examine a small handful of style types over time therefore lack the ability to make comparative statements about the *relative* magnitude of gendered differences. I contribute by measuring a range of styles that are congruent with both masculine agentic stereotypes and feminine communal stereotypes. Finally, most studies have examined gendered differences at a fixed point in time. This seems an important oversight, as gender role theorists contend that stereotypes are dynamic concepts and that the content of stereotypes associated with men and women may change over time as the distribution of gendered societal positions change (Diekmann and Eagly, 2000; Eagly and Wood, 2012; Eagly et al., 2020). As the content and validity of stereotypes have changed over time, then it seems likely that the sets of behaviours that men and women are expected to express will also have changed. I provide an important contribution to existing work by being the first to measure the extent to which politicians' conformity to gender stereotypes has changed over time.

## 2.4 Policy-based behavioural differences

Just as stereotypes are said to inform the styles of men and women in political office, so too may they inform the policy areas that men and women are expected to focus on. Building on the work described above, Eagly (2013) argues that differences in the political attitudes and policy priorities of men and women emerge from socialisation and gender role attitudes. In reviewing explanations for gender differences, Huddy, Cassese and Lizotte (2008) suggest that differences in political attitudes are related to the personality traits, such as agreeableness, empathy, and assertiveness, that men and women are said to have. Women are said to orientate towards policies that focus on caring for others (Beutel and Marinin, 1995) and "compassion" issues such as poverty, education, children, and health policy issues (Shapiro, 1986).

Work from a range of countries and contexts has shown that there is strong empirical evidence that men and women politicians differ markedly in the policies and issues they prioritise when in office. These differences can manifest in the sets of policies politicians support (Piscopo, 2011; Swers,

2002; Thomas, 1994), the debates they participate in (Bäck, Debus and Müller, 2014; Bäck and Debus, 2019), the committees and leadership roles they hold (Barnes and O'Brien, 2018; Escobar-Lemmon and Taylor-Robinson, 2005; Goodwin, Holden Bates and McKay, 2021; Krook and O'Brien, 2012), speeches they deliver (Clayton, Josefsson and Wang, 2017; Dietrich, Hayes and O'Brien, 2019), and in the questions they ask (Bird, 2005). In interviews with legislators, too, men and women report different policy priorities (Schwindt-Bayer, 2006; Skjeie, 1991; Wängnerud, 2000).

There is ample debate in the literature about the sets of issues we should expect men and women to champion, or whether it is appropriate to categorise *a priori* particular issues as associated with men or women (Beckwith, 2014; Celis et al., 2014). Commonly, a distinction is made between social policy issues such as health care, welfare, education, family, and culture that relate to women's traditional role as caregivers (Norris, 1996), and the long-standing stereotype that women are more communal, caring, and compassionate (Eagly and Karau, 2002), and issues such as economic affairs, defence, employment, and trade as associated with men (Krook and O'Brien, 2012). Other work has extended this to the focus on egalitarian policies more broadly (Campbell and Childs, 2010; Wängnerud, 2000), and studies suggest that increasing the numbers of women in parliament has a positive impact on the degree to which legislation focuses on issues such as child care, family, or equal wage policies (Annesley, Engeli and Gains, 2015; Bratton and Ray, 2002), that the introduction of gender quota laws, which are associated with the increased numerical presence of women, may also lead to greater focus on social justice issues (Weeks, 2019), as may women's presence in parties more broadly (Greene and O'Brien, 2016; Kittilson, 2011).

In the UK context, Catalano (2009) assesses, first, whether there are gender differences in the extent to which politicians participate in debates on the two issue areas and, second, whether a difference indicates that women politicians use parliamentary debates to achieve greater substantive representation of women's interests. She finds that women are significantly more likely than men to participate in debates on health care topics, even when controlling for factors such as party, seniority, or relevant committee membership, however, finds no significant effect of gender on participation in finance related debates. Focusing instead on politicians' use of parliamentary questions, Bird (2005) assesses whether there are gender differences in the extent to which MPs ask questions related to women, men, or gender more broadly. She finds that women are significantly more likely than men to refer to "women" and "gender" in both oral and written questions. Further, while overall she uncovers intuitive differences that align with expectations from gender role theory in the departments where questions related to women were asked (Social Security received the most and the Treasury among

the least), women politicians questioned a wider range of departments on issues related to women and gender than men.

Although not as developed as the literature on measuring elite behaviour, other work has suggested these differences may also be present in the priorities of voters too, where, relative to men, women have been shown to be more concerned with “compassion” issues (Page and Shapiro, 2010; Schneider and Bos, 2019) and have taken positions with a more compassionate stance, such as being less supportive of military action (Fite, Genest and Wilcox, 1990; Togeby, 1994), and more supportive of preventing crime (Hurwitz and Smithey, 1998). Further, women are less supportive of the use of force than men, and more supportive of welfare policies and government spending (Campbell, 2012). In recent work, Yildirim (2021) adopts an inductive approach to identifying the sets of issues men and women prioritise and finds that women show particular concern for policy areas traditionally associated as “women’s issues”, such as health policy, education, housing, poverty, family and children, welfare, and crime, whereas men are more likely to prioritise issues concerning foreign trade, economic and foreign policies, farming, energy, taxation, corruption, and immigration. Other work has suggested that politicians and parties are aware of these differences in the preferences of women in the public, and that this informs strategies in campaigns and in legislative behaviour at both the individual-level (Ennsner-Jedenastik, 2017; Herrnson, Lay and Stokes, 2003; Höhmann, 2020) and the party-level (Campbell and Childs, 2015; Childs and Webb, 2012; Meguid et al., 2022).

Overall, this broad review of the literature on differences in the policies and issues men and women prioritise suggests that divisions are broadly in line with traditional stereotypical expectations. In chapter 4, I focus on the sets of issues politicians in the UK raise and make several contributions to the literature described above. First, while evidence from politicians’ testimonies in the UK suggest men and women MPs may differ in the sets of issues they raise (Childs, 2004b), with few exceptions (Bird, 2005; Catalano, 2009), there is currently only limited systematic evidence that this is the case in the UK. I make progress by addressing these questions at scale, both with respect to the time-horizon of the study and number of policy areas studied.

Second, I examine not only whether there are gender differences at the aggregate level, but also consider how these differences may change across political careers. At the beginning of their careers, politicians are at their most inexperienced and will therefore gravitate towards the set of policy areas where they both have the most prior experience (Bailer and Ohmura, 2018; Bailer et al., 2022) and, crucially, will be *ascribed* the most authority and competence in by others to gain credibility and competence in their roles. For women, because of both their shared identity with women in the popu-



lation (Wängnerud, 2009) and the argument that they are better suited to and have greater expertise in traditionally “feminine” issue areas (Herrnson, Lay and Stokes, 2003), they are likely to emphasise traditionally feminine policies. However, as politicians spend more time in office, they learn to navigate their roles, and establish themselves as experienced in a wider variety of policy areas (Miquel and Snyder, 2006; Rush and Giddings, 2011), the selection of issues they raise will diversify beyond those stereotypically associated with their gender. My central argument is that women politicians will focus more on traditional women’s issues than men, but that this gap will be largest among junior politicians and will diminish with increased experience.

## 2.5 Stereotypes and voters’ perceptions and evaluations

Do voters punish politicians who behave in a way that is incongruent with gender-role expectations? I distinguish between how past work has studied the content of stereotypes, and between the ways in which individuals may apply gender stereotypes when judging men and women’s behaviour. While existing literature, described below, is inconclusive on the exact influence gender stereotypes have on voting outcomes, prior work presents clear distinctions with respect to the content of voters’ gender stereotypes. First, voters assume that women are more competent with respect to “feminine” or “women’s” issues including education, health care, and welfare, whereas men are seen to be stronger on issue such as the economy, policing, or the military (Huddy and Terkildsen, 1993a; Kahn, 1996; Lawless, 2004; McDermott, 1997). Second, women are often viewed as more compassionate and communal, and that men are viewed as more tough and agentic (Brooks, 2013; Dolan, 2004; Hayes, 2011).

There are two ways in which individuals may apply gender stereotypes when judging the behaviour of men and women. The first, which characterises most studies in political science, is the extent to which voters differentially *evaluate* men and women based on stereotypical expectations (Anzia and Bernhard, 2022; de Geus et al., 2021; Saha and Weeks, 2022). These studies address questions such as whether gender stereotype-(in)congruent behaviours lead voters to evaluate politicians to be more or less competent, likeable, charismatic, electable, hard-working, or honest based on their gender. Second, individuals may differentially *perceive* men and women’s behaviour based on stereotypical expectations (Biernat, Manis and Nelson, 1991; Biernat and Manis, 1994; Koch, 2000). Evaluation-based differences focus on how politicians’ use of particular behaviours, say, expressing anger, leads to voters evaluate them as more or less electable. Instead, perception-based differences focus on whether voters actually just perceive anger to be different if expressed by a man or a woman.



### *Stereotypes and behavioural evaluations*

Huddy and Terkildsen's (1993a) seminal early work in the US has been crucial in forming experimental studies on gender stereotypes. In the paper, they investigate the origins of voters' expectations that women politicians should be more competent on "compassion" based issues, such as health care or welfare, and men's supposed greater competency on financial, military, and defence issues. They make the distinction, first, between *gender-trait* based stereotypes, emphasising the relationship between a politician's gender-linked personality traits, where women are believed to be "warm, gentle, kind and passive", where men are perceived as more "tough, aggressive and assertive" (Huddy and Terkildsen, 1993a, 121). Second, they outline *gender-belief* based stereotypes, which instead focus on the political outlooks of men and women, such as the idea that women have more of a left-wing ideological outlook than men. In the experiment, they find strong support for the trait-based approach: candidates stereotyped with warm and expressive traits were evaluated as better able to handle compassion-based issues, such as health, welfare, and poverty, whereas candidates stereotyped with masculine traits were evaluated as more competent on issues such as the military and defence, but also on a wider range of political issues.

The findings from Huddy and Terkildsen (1993a) have been instrumental in informing subsequent studies on stereotypes and issue competency. For instance, in a variety of studies, Holman, Merolla and Zechmeister emphasise the importance of the policy domain in which politicians operate for the influence of stereotypes (2011; 2016; 2017). They find that when the policy domain focuses on a stereotypically masculine issue area, here national security and terrorist threat, women candidates from the Democratic Party are negatively evaluated compared to women from the Republican Party and men from both parties (Holman, Merolla and Zechmeister, 2011, 2016, 2017). Other work supports this finding, showing that women candidates in the US are most successful when they run "as women", by emphasising women's issues, such as compassion issues, traditional values, and traditional women's issues (Herrnson, Lay and Stokes, 2003). Other work has suggested this may feed into the type of elected office, where women candidates will be disadvantaged least in elections to office that are congruent with women's issue areas. Among other things, Anzia and Bernhard (2022) compare women's success in running for school board elections compared to city council races and find that women have more success in races for elected office when the salient issue is stereotype-congruent.

Further studies have sought to assess how stereotypes about the policy areas men and women are expected to have authority influence how persuasive voters find arguments to be. Anderson-Nilsson

and Clayton (2021) argue that men will be most persuasive when delivering arguments in neutral or “masculine” policy areas, however women will be evaluated as more persuasive in “feminine” policy areas or on policies related to women’s rights specifically. However, they find that neutral policies are more persuasive when delivered by women, and that “feminine” and “masculine” policy areas are equally persuasive when delivered by men and women. In the context of political advertising, Searles et al. (2020) argue that men are most effective when advertising masculine issues and women when advertising feminine issues. However, they find no evidence of differential gender effects. Overall, therefore, work suggests that the policy domain in question seems to influence how voters evaluate politicians. As such, women politicians may be incentivised to focus more on stereotypically “feminine” policy areas, while avoiding traditionally “masculine” policy domains.

The second way in which stereotypes inform voters’ evaluations of politicians is with respect to the styles or behaviours they express in office (Brooks, 2013; Dolan, 2004; Hayes, 2011). As outlined in section 2.1, a key distinction is drawn between communal traits which are associated with women and “masculine” traits which are associated with men. A wealth of studies has sought to assess whether politicians are rewarded when they behave in a way that is congruent with the stereotypes associated with their gender roles, or whether they are punished when they instead behave in a way that violates these gender roles. Voters are said to prefer leaders who express gender-role-congruent emotions (Fischbach, Lichtenhaler and Horstmann, 2015; Klofstad, Anderson and Nowicki, 2015), and recent observational work has shown that women politicians are punished for violating stereotypical emotions. Leveraging a variety of video, audio, and text techniques, Boussalis et al. (2021) focus on the behaviour of Angela Merkel and find that she is rewarded by voters when she displays happiness – a style consistent with feminine communal stereotypes – but punished for displays of anger – a style consistent with masculine agentic stereotypes.

Many other studies, however, have suggested that stereotypes *overall* may not inform voters’ evaluations, but that stereotype application depends on the context or type of voter. For instance, Bauer (2015a) shows that stereotypes around emotionality do harm women, but only when they are explicitly activated. Similarly, Krupnikov and Bauer (2014) and Cassese and Holman (2018) find that women are punished for negativity, a stereotype-incongruent style, although only by voters from a different political party. Similarly, Bauer, Kalmoe and Russell (2022) find that voters dislike uncivil candidates overall, but that women incur no particular punishment except by voters who themselves have aggressive personalities. In other work, Bauer (2015b) has shown that not all voters stereotype politicians, but that a voter’s partisanship, political knowledge, and political attention matter. In contrast, Brooks

(2011) finds no evidence of gender bias in voters' evaluations of candidates, but rather that men and women are similarly penalised for displays of anger and tears during elections. Recent work consistently emphasises that the power of stereotypes depends on the electoral context, stereotype under scrutiny, or on the type of voter.

Stereotypes around competence, which are related to agentic stereotypes (Cuddy, Fiske and Glick, 2008), have also been shown to be more compatible with men. Studies focusing on these characteristics have shown that voters require more reassurance about the competence of women compared to men (Ditonto, Hamilton and Redlawsk, 2014; Ditonto, 2017), and that women candidates therefore face higher standards when they run for office and require better qualifications in order to overcome stereotypes (Bauer, 2020a,b). Although other studies from the UK suggest that women may not face higher standards than men (Blumenau and Lauderdale, 2022a). Eggers, Vivyan and Wagner (2017) assess whether there are gender differences in the punishments politicians incur for corrupt behaviour in office and uncover no evidence of gender differences in the punishment afforded for misconduct. Similarly, in recent work has assessed whether women must work harder in office to be evaluated the same as men (Hargrave and Smith, 2022), and find that voters overall prefer politicians who are productive, and reward productive politicians in job performance and electability evaluations. However, they find no evidence that productive women are unjustly rewarded, nor do unproductive women face greater punishment than men.

Related to competence is the idea of being tough and powerful in office, and recent experimental work by Utych (2021) has shown that voters prefer candidates who are powerful as opposed to powerless in campaigns, although women are not differentially penalised compared to men. Relatedly, Saha and Weeks (2022) assess whether ambitious women are punished in the UK and the US, arguing that women who are ambitious are likely to face backlash. They find that ambitious women in both contexts are not penalised overall compared to men, and that while in the US there are noticeable partisan differences, this is not the case in the UK.

The influence that gender stereotypes have on voters' evaluations of politicians, therefore, seems to vary according to the type of campaign or voter, or the stereotype studied. However, as I argue in further detail in chapter 5, another explanation for varying voter evaluations may be that stereotypes lead voters to differentially *perceive* the behaviour of men and women. While this argument has received very limited attention in the political science literature, this question is of central concern in social psychology work.

*Stereotypes and behavioural perceptions*

A central argument in the social psychology literature is that individuals will regularly shift and adjust their perceptions as they judge members of different social groups, and that these judgements will be informed by stereotypical priors. The idea here is that when judging individuals from different social groups, one may change their perceptions based on their expectations of the pre-existing standard for that group. A common example of subjective perceptions provided by prior work in social psychology is perceptions of an individual's height (Biernat and Manis, 1994). If an individual is asked whether they perceive 5 foot and 9 inches woman to be tall, they will likely base this assessment on what they perceive to be the average height for women – which, for instance, is 5 foot and 3 inches in the UK.<sup>1</sup> Since 5 foot and 9 inches is significantly above the average height for a woman, the woman is likely to be perceived to be tall. However, if an individual is asked instead whether they perceive a man who is 5 foot and 9 inches to be tall, they will likely use the average height for men – which, for example, is 5 foot and 9 inches in the UK – as the benchmark. Since 5 foot and 9 inches matches men's average height, they may not perceive the man to be particularly tall. In this example, the categorisation of an individual as “tall” will differ depending on our prior expectations about a given group. As such, our perceptions of the qualities of individuals are subjective and depend on the pre-existing, often stereotypical, expectations that we hold for members of a given group

Many social psychology studies have shown that these biased perceptions, which can occur both consciously or subconsciously, apply when assessing men and women with respect to traits (Blair and Banaji, 1996), professions (Banaji and Greenwald, 1995; Biernat, Manis and Nelson, 1991), or writing styles (Strand, 1999; Francis et al., 2003; Knobloch-Westerwick, Glynn and Huges, 2013). With respect to behaviour specifically, Biernat and Manis (1994) assess whether there are different thresholds for categorising men and women as aggressive, assertive, or unassertive. They argue that because of prior stereotypical expectations that men are perceived to be more aggressive, participants will have a higher threshold for labelling behaviour as aggressive when committed by men. However, as aggression is gender role-incongruent for women, the threshold for them to be labelled as “very aggressive” would be more easily met. Similarly, as women are expected to be unassertive and men are expected instead to be assertive, the threshold for men to be labelled as “very unassertive” would be lower than for women. To assess this expectation, Biernat and Manis (1994) presented respondents with lists of behaviours representative of aggression, unassertiveness, or assertiveness to see how many of the

---

<sup>1</sup>See “Statistics reveal Britain's ‘Mr and Mrs Average’”, BBC News, 13th October 2010.

behaviours men and women needed to meet to be classified as aggressive, unassertive, or assertive. For instance, examples of aggressive behaviour included “grabbed his/her wallet back from a pick-pocket on the bus”, while “was talked into going to see a fairly bad movie for a second time” was an example of an unassertiveness behaviour, and “drew up a petition and persuaded people to sign it” for assertiveness. Their findings support their expectations for aggression and unassertiveness, but not for assertiveness.

Banaji, Hardin and Rothman (1993) conduct a similar study, however, unlike Biernat and Manis (1994), argue that individuals will be stereotyped to be *more* like the stereotypical categories they are associated with. Specifically, they argue that because women are expected to be dependent, they will be judged as such, and because men are expected to be aggressive, they will, compared to women, be judged as such. They find no evidence that women are perceived to be more dependent, but they find that exposure to aggression influences perceptions of men but not women.

Social psychology studies suggest that the application of stereotypes in perceptions of men and women's behaviour is common. Few political science studies, however, have assessed whether there are differences in voters' perceptions of men and women. Indeed, to the best of my knowledge, this has thus far been studied only with respect to perceptions of politicians' ideological leanings. In several studies, Koch assesses voters' perceptions of how liberal men and women are (2000; 2002), arguing that voters will perceive women candidates in US Congressional elections to be more liberal than men. Using data from the 1988, 1990, and 1992 Pooled Senate Election Study, Koch (2000) finds strong evidence that voters apply gender stereotypes to infer candidates' ideological positions. Women candidates running for the Democratic Party are perceived as more liberal than they actually are, and women candidates running for the Republican Party are instead perceived as less conservative than they actually are.

Many studies in social psychology (Banaji and Greenwald, 1995; Banaji and Hardin, 1996; Banaji, Hardin and Rothman, 1993; Biernat, Manis and Nelson, 1991; Biernat and Manis, 1994) and very few studies in political science (Koch, 2000, 2002) have shown that prior gender stereotypical expectations inform how men and women's behaviour is perceived. While social psychology suggests that stereotypes do inform voters' perceptions, the evidence for the political context is limited to only work on partisanship from nearly 20 years ago. The current scholarly evidence on whether voters perceive men and women candidates' or politicians' behaviour to be different is limited. In chapter 5, I provide an important contribution to this literature by addressing whether stereotypes inform both voters' evaluations *and* perceptions of politicians.

## 2.6 Conclusion

I conclude this chapter by outlining in brief the theoretical and empirical contribution that each of the subsequent chapters in this thesis strives to provide. First, a common motivation outlining why men and women may express different political styles are they may internalise this expectation prior to entry office and anticipate backlash from voters for behaviour that violates stereotypical norms. The stereotypes associated with men and women arise initially from men and women's historical occupation of different social roles which are associated with different characteristics. Gender role theorists contend that stereotypes are dynamic concepts and that the content of stereotypes associated with men and women may change over time as the distribution of gendered societal positions have changed (Diekman and Eagly, 2000; Eagly and Wood, 2012; Eagly et al., 2020). Therefore, if the content and validity of stereotypes have changed, then the sets of behaviours that men and women are expected to express will also have changed. Despite efforts from social psychologists to emphasise the *dynamic* nature of stereotypes, political scientists have neglected to assess how gendered patterns of behavioural styles may vary over time.

In chapter 3, I build on the work on dynamic stereotypes, and argue that – as a result of changes in the public's conformity to (Twenge, 2001; Leaper and Ayres, 2007) and perceptions of the validity of gender stereotypes (Taylor and Scott, 2018; Sendén et al., 2019; Shorrocks, 2018) and increases in the presence and prominence of women in parliament (Blumenau, 2021c; Dasgupta and Asgari, 2004; Diekman et al., 2005) – incentives to conform to gender-role consistent behavioural stereotypes are likely to have weakened in recent years. As a consequence, these arguments lead to a central behavioural prediction which I test empirically: that UK MPs will conform less to gender stereotypes now than in the past, and that women in particular will adopt styles that are further from feminine stereotypes over time.

Second, stereotypes are also said to guide the sets of issues that politicians will raise and prioritise during their time in office. Thus far, existing perspectives have shown that, although context may mitigate the link (Childs and Krook, 2009; Hawkesworth, 2003), women politicians overall have been shown to raise issues that are traditionally associated with women and enhance women's experiences in the policy process more broadly, and that this is observable in various aspects of their legislative behaviour. However, despite the growing work that has emphasised how incentives shape politicians' and parties' representational efforts (Bailer et al., 2022; Weeks et al., 2022; Höhmann, 2020; Höhmann and Nugent, 2021; Meguid et al., 2022; Valdini, 2019), existing work has largely not considered

how these dynamics may change over the course of politicians' careers. As women have historically been underrepresented numerically in legislatures and have, on average, less experience in office than men (Lazarus, Steigerwalt and Clark, 2022), it has been challenging for studies to date to identify the implications that increased political experience may have for both the behaviour of women legislators in office and for the representation of voters' interests in the policy-making process. In chapter 4, I seek to understand the effect of politicians' increased experience on their representational behaviour. I argue that gender differences in the extent to which women politicians raise issues traditionally associated with women will be most pronounced among junior politicians who lack experience and credibility in the eyes of voters and their colleagues, however, this difference will decrease with increased parliamentary prestige. In doing so, I provide important insights into the development of gendered behaviour in parliament going forward as the numbers and experience of women politicians may continue to increase.

Finally, stereotypes are also said to wield significant influence on how voters engage with politicians' behaviour. While there is a robust literature, described above, on the influence of gender stereotypes on voter evaluations in politics more broadly (Bauer, 2020b; Hargrave and Smith, 2022; Saha and Weeks, 2022), in chapter 5, I contribute to this literature in several key ways. First, work to date has tended to focus on how voters evaluate stereotype-congruent or incongruent behaviour by focusing on isolated behaviour traits, such as tears (Brooks, 2011), ambition (Saha and Weeks, 2022), negative attacks (Krupnikov and Bauer, 2014), or willingness to compromise (Bauer, Harbridge Yong and Krupnikov, 2017). Instead, I expand upon this to focus on how voters evaluate politicians' use of a wider range of styles that are consistent with both feminine communal and masculine agentic stereotypes. In doing so, I can assess whether voters equally penalise men and women for stereotype-incongruent behaviours. Second, despite the rich literature on voter evaluations in politics more broadly, so far political scientists have not assessed whether voters differentially *perceive* politicians' behaviour based on gender alone, and whether differential perception of behaviour itself may in turn be responsible for how voters evaluate politicians. That is, even in the absence of any objective differences in, say, how aggressive a politician is, voters may perceive the politician to be more or less aggressive depending on whether they are a man or a woman. This differential perception may, in turn, lead to voters' differential evaluation of politicians. In chapter 5, I make an important contribution to past work by being the first study to assess whether differential perceptions of styles themselves may serve as a potential mechanism through which voters' gendered evaluations of politicians might become manifest.

Taken together, the empirical chapters in this thesis strive to contribute to the rich body of existing

work reviewed here by emphasising the dynamic power of stereotypes in influencing how politicians behave and how voters may engage with this behaviour.



## Chapter 3

# No Longer Conforming to Stereotypes? Gender, Political Style, and Parliamentary Debate in the UK

A version of this chapter is published as: Hargrave, Lotte and Jack Blumenau (2022) “No Longer Conforming to Stereotypes? Gender, Political Science, and Parliamentary Debate in the UK”, *British Journal of Political Science*, 52(4): 1584–1601. <https://doi.org/10.1017/S0007123421000648>.

### Abstract

Research on political style suggests that where women make arguments that are more emotional, empathetic, and positive, men use language that is more analytical, aggressive, and complex. However, existing work does not consider how gendered patterns of style vary over time. Focusing on the UK, I argue that pressures for women politicians to conform to stereotypically “feminine” styles have diminished in recent years. To test this argument, I describe novel quantitative text analysis approaches for measuring a diverse set of styles at scale in political speech data. Analysing UK parliamentary debates between 1997 and 2019, I show that women MPs’ debating styles have changed substantially over time, as women in parliament have increasingly adopted stylistic traits that are typically associated with “masculine” stereotypes of communication. These findings imply that prominent gender-based stereotypes of politicians’ behaviour are significantly worse descriptors of empirical reality now than they were in the past.

### 3.1 Introduction

Have incentives for politicians to conform to gender stereotypes diminished over time? In addition to the fact that men and women politicians speak about systematically different sets of political issues (Bäck and Debus, 2019; Catalano, 2009, and addressed in chapter 4), another dimension on which gendered differences are said to arise is regarding argumentation style. Gendered communication styles are thought to be rooted in stereotypes that create social expectations for women to act “like women” and men “like men” (Eagly and Wood, 2012). If politicians internalise these expectations before entering politics, or if voters punish them for contravening gender stereotypes (Bauer, 2015a; Boussalis et al., 2021; Cassese and Holman, 2018), legislators are likely to engage in gender-role consistent behaviour, and we should expect systematic differences in the political styles that men and women politicians adopt. Empirical evidence supports this view: compared to men, women politicians’ speeches are more emotional (Dietrich, Hayes and O’Brien, 2019), less complex and jargonistic (Coates, 2015), less repetitive (Childs, 2004b), less aggressive (Kathlene, 1994), and use different types of evidence to support their arguments (Hargrave and Langengen, 2021).

I contribute to this literature by evaluating the degree to which gendered differences in political style vary over time. Consistent with work that views stereotypes as dynamic constructs (Eagly and Wood, 2012; Diekman and Eagly, 2000), I argue that over the past 25 years in the UK, where I situate the study, several factors are likely to have decreased the degree to which UK MPs (and especially women) will conform to stereotype-consistent behaviours in parliament. First, politicians are drawn from a broader population which has itself diverged from stereotypical communicative styles in recent years. Second, changes to the social roles played by women in public life, and in politics, have reduced the validity of gender stereotypes in the eyes of the public. Consequently, I argue that voters are less likely to sanction women legislators for gender-incongruent behaviours now than in the past. Finally, the increased prominence of women in parliament and leadership roles is also likely to reduce the degree to which women politicians internalise expectations that they need to behave in “feminised” ways. Together, these arguments lead to a central behavioural prediction which I test empirically: that UK MPs will conform less to gender stereotypes now than in the past, and that women in particular will adopt styles that are further from feminine stereotypes over time.

To evaluate this expectation, I examine politicians’ styles as they manifest in one prominent legislative activity: parliamentary debates. I conceive of *debating* style as a characteristic of speech which is distinct from its content. Intuitively, the style of a speech reflects the manner in which an argument

is delivered. In social psychology, women's communal styles are thought to be marked by higher levels of emotionality, positivity, empathy, and warmth, while men's agentic styles are thought to be marked by higher levels of aggression, logic, and confidence (Eagly and Wood, 2012; Schneider and Bos, 2019). In political science, these concepts have been operationalised using a diverse set of indicators. I survey both literatures and identify eight styles that reflect the ideas of communality and agency, and are also – in principle – detectable in the speeches politicians deliver. The eight styles which I use as the basis of my empirical analysis are human narrative, affect, positive emotion, negative emotion, factual language, aggression, complex language, and repetition.

In addition to my substantive argument, this chapter also provides a methodological contribution to the measurement of style in legislative settings (Boussalis et al., 2021; Dietrich, Hayes and O'Brien, 2019). My goal is to construct measures that closely approximate the conceptual definitions of each of the styles that I highlight in the review of the literature. For some styles, I use existing quantitative text analysis measures that have been extensively validated in other settings (e.g. Kincaid et al., 1975). For others, I develop new measures that combine traditional dictionary approaches with a word-embedding model. My strategy overcomes limitations of standard dictionary approaches as it enables me to detect styles as they manifest *in the specific context of parliamentary debate*. Evidence from a human validation task shows that these measures significantly outperform standard measures that have been used extensively in previous research on political style.

I apply the measures to nearly half a million speeches delivered in the House of Commons between 1997 and 2019 and report three main findings. First, in the early parts of the study period, I document patterns of style which are broadly consistent with expectations from the literature on gender stereotypes. Men's speeches are marked by substantially higher levels of aggression and complexity, while women MPs' speeches are considerably more emotional, positive, and make greater use of human narrative. Second, and crucially, I find that these differences have reduced dramatically in recent years. In six out of eight styles I study, MP gender explains less variation in style use between individuals at the end of the study period than at the beginning. The magnitudes of these changes are large: the proportion of MP-level variation explained by gender decreased by between 29% and 57% from the beginning to the end of the time period, depending on the style. For cases where I report *diverging* behaviour over time, these stylistic shifts run counter to gender-role expectations. Third, I show that the evolving variation in style use has primarily resulted from women's decreasing use of communal styles, and increasing use of agentic styles, over time.

My work builds upon a large literature on gender differences in politicians' communication styles,

the vast majority of which considers whether gender stereotypes accurately capture real behavioural differences at fixed points in time. By contrast, I show that the descriptive validity of prominent stereotypes of how men and women communicate is considerably lower in the contemporary House of Commons than it was in the past.

These findings also have important implications beyond scholarly accounts of legislative politics. Prescriptive stereotypes for how men and women *ought* to behave are ultimately rooted in our collective understanding of how we expect men and women *will* behave (Eagly and Wood, 2012; Gill, 2004). In politics, these prescriptions can form the basis of voter judgements about the behaviour of men and women politicians. Documenting when behavioural shifts run counter to gender-based stereotypes is important, then, because it potentially undermines these prescriptions, and could thereby diminish the degree to which women will be subject to penalties for failing to conform to stereotypical expectations (Cassese and Holman, 2018; Ditonto, 2017).

Additionally, these results cast doubt on the idea that the election of more women into office will automatically result in a less adversarial and more deliberative culture in Westminster.<sup>1</sup> At least in the context of parliamentary debate, my findings suggest that such effects are unlikely to materialise because they are based on an outdated assumption about the distinctiveness of women MPs' political styles. In addition, by placing hopes of cultural change on newly elected women, proponents of these views may also be setting women politicians up to fail if their increased presence does not result in a "better" political culture. To the extent that cultural change of this sort is a desideratum of modern parliamentary politics, these results suggest that hopes of affecting such change should not rely on the presumption that newly elected women will conform to anachronistic stereotypes, and that purposive reforms to parliamentary practices may instead be necessary.

### 3.2 Gender, stereotypes, and debating style

Why might men and women in parliament employ different debating styles? Gender role theory (Eagly and Karau, 2002) suggests that gender stereotypes concerning the typical conduct of men and women can affect behaviour via two main channels. First, repeated exposure to stereotypes from a young age lead men and women to *internalise* expectations relevant to their genders, which then become self-imposed standards against which they regulate their own behaviour (Eagly and Wood, 2012). Second, descriptive stereotypes (e.g., the perceived tendency for women to be emotional) are often thought

---

<sup>1</sup>See, for example, [Designing a new parliament with women in mind](#), *Democratic Audit*, 29th July 2016.

to lead to prescriptive stereotypes (e.g., the view that women *should* be emotional), the violation of which leads to the imposition of *social sanctions* by others which further incentivise conformity with gender-based norms (Brescoll and Uhlmann, 2008).

Women politicians are especially subject to pressures to conform to role-consistent behavioural standards, as voters punish women for displaying behaviour that counters feminine stereotypes. Voters form gender-biased impressions of candidates (Bauer, 2015a), and penalise women for appearing to be too ambitious (Okimoto and Brescoll, 2010) or negative (Cassese and Holman, 2018), while rewarding them for displays of happiness (Boussalis et al., 2021). These penalties are more acute when the campaign environment is characterised by “masculine” issues (Holman, Merolla and Zechmeister, 2016), and are more commonly applied by low-attention voters (Bauer, 2015b) and voters with sexist attitudes (Mo, 2015) or aggressive personalities (Bauer, Kalmoe and Russell, 2022). By contrast, voters are less sensitive to role-inconsistent behaviour by men (Okimoto and Brescoll, 2010).

However, as political leaders, legislators are also expected to display behaviours consistent with *leadership* stereotypes. Men’s historical occupation of leadership positions means that leadership stereotypes have been shaped by traditional “masculine” traits such as being assertive, competitive, and outgoing (Koenig et al., 2011). The congruence between leadership stereotypes and masculine stereotypes therefore poses little challenge for men in office to conform to both sets of expectations. By contrast, if women seek to conform to *leadership* stereotypes, they risk incurring penalties for violating *feminine* stereotypes (Bauer, 2017; Eagly and Karau, 2002). Women must therefore attempt to balance a complicated array of behavioural expectations in a way that men do not.

Political scientists have evaluated whether these incentives induce men and women politicians to adopt systematically different political styles. I focus on one aspect of legislative activity where differences are likely to become manifest: in political speech. On which dimensions of style should we expect gender differences? Of central concern in the social psychology literature is the distinction between *communal* characteristics of style, which are associated with women, and *agentic* characteristics which are associated with men (Schneider and Bos, 2019). These labels are heuristics for clusters of behavioural attributes, where communal characteristics are said to include being “affectionate, helpful, kind, sympathetic, interpersonally sensitive, nurturant, and gentle”, while agentic characteristics include being “aggressive, ambitious, dominant, forceful, independent, self-sufficient, and self-confident” (Eagly and Karau, 2002, 574). By surveying the large empirical literature on gendered styles in political science, I identified eight styles that are representative of communal or agentic behaviour and which previous work had either shown to be associated with men and women use in politics, or

that previous work had *expected* to be associated with gender differences. I use these styles as the basis of the empirical analysis below.

I identified three communal characteristics of style that are typically associated with women. First, women are said to make greater use of **human narrative** through reliance on personal experience, analogies, and anecdotes in their speeches (Blankenship and Robson, 1995). This idea is supported both by politicians' testimonies (Childs, 2004b), and qualitative studies of political speech (Hargrave and Langengen, 2021). Second, women are also thought to make greater use of emotional language or **affect** (Huddy and Terkildsen, 1993a), and there is clear evidence that women's language exhibits greater overall emotionality than men's (Dietrich, Hayes and O'Brien, 2019; Jones, 2016). Third, and more specifically, women have been found to use more **positive emotion**, such as expressing happiness, in their political speeches than men (Boussalis et al., 2021; Yu, 2013).

I identified five agentic characteristics of style that are typically associated with men. First, men are thought to rely more on **fact-based** language, which is more "analytical, organised and impersonal" and relies more on statistical evidence (Jamieson, 1995, 76). In the UK, MPs suggest that men pay greater attention to "scientific research" (Childs, 2004b, 181), though there is evidence that women use more factual language (Hargrave and Langengen, 2021). Second, men's speech is also thought to feature higher levels of linguistic **complexity**, marked by formalistic and jargonistic word use (Childs, 2004b), while women are thought to be more accessible and clear (Coates, 2015). Third, men are also thought to be more **repetitive** (Dahlerup, 1988; Childs, 2004b, 184). Fourth, men are said to be more **aggressive**, whereas women are said to avoid combative and aggressive styles (Brescoll and Uhlmann, 2008; Kathlene, 1994), and empirical work suggests women are significantly less adversarial than men in parliamentary debate (Grey, 2002; Hargrave and Langengen, 2021). Fifth, women are thought to avoid the use of excess **negative emotion** for fear of backlash (Cassese and Holman, 2018), while men are thought to make greater use of negativity (Brooks, 2011).

I summarise these eight styles in table 3.1. I provide a short definition and categorise each as either communal or agentic according to the discussion above. The expectations that derive from the literature on gendered stereotypes suggests that women will be more likely to use communal debating styles, and men more likely to use agentic debating styles.

Table 3.1: Political styles

Style	Type	Definition
Human Narrative	Communal	Use of personal examples or experiences; stories of other people; constituency stories; illustrative examples; referring to individuals.
Affect	Communal	Use of emotive language, which might be either positive or negative; such as expressing criticism, praise, disapproval, pride, empathy or fear.
Positive Emotion	Communal	Use of positive emotional language, which might include expressing empathy, praise, celebration or congratulations.
Fact	Agentic	Use of numbers, statistics, numerical quantifiers, figures and empirical evidence.
Complexity	Agentic	Use of jargonistic, complicated and elaborate language that is challenging to understand.
Repetition	Agentic	Repeated use of the same words or phrases.
Aggression	Agentic	Use of aggressive or combative language, which might include criticisms or insults; language that suggests forceful action; or declamatory or adversarial language.
Negative Emotion	Agentic	Use of negative emotional language, which might include expressing fear, anxiety, unpleasantness, sadness or disapproval.

### *Dynamic gender stereotypes*

Despite this rich literature, few studies consider whether politicians' conformity with gender stereotypes has changed over time.<sup>2</sup> This is surprising, as gender role theorists emphasise that the content and strength of stereotypes are dynamic (Diekmann and Eagly, 2000; Eagly and Wood, 2012). These accounts posit that gender stereotypes arise from men and women's historical occupation of different social roles which are associated with different characteristics. For instance, because women have traditionally occupied roles in which they provide care to others, "caring" as a characteristic became stereotypic of women. However, as the distribution of men and women into different roles changes, so too will the characteristics associated with the stereotypes themselves such that the stereotype of women will be marked by "increasing masculinity and... decreasing femininity" (Diekmann and Eagly, 2000, 1173). Building on this logic, I argue that recent changes in the roles played by women in both politics and the broader public are likely to have weakened traditional gender stereotypes in the UK, and I therefore expect a decline in the degree to which MPs, and especially women, will adopt styles that are congruent with the stereotypes described above.

First, politicians are selected from a broader population, which has itself diverged from gender-

<sup>2</sup>Though see Jones (2016) for a case study of the evolution of Hillary Clinton's style, and Grey (2002) who demonstrates that women MPs in New Zealand are increasingly aggressive over time.

stereotypical behaviours over time. In most advanced economies in recent decades, women's traditional role as care-givers has declined, and women's educational attainment, participation in the workforce, and occupancy of senior management positions have increased (Sayer, Bianchi and Robinson, 2004; Diekman and Goodfriend, 2006, 370). As societal gender roles have changed, women in the public have come to demonstrate increasingly agentic behaviours across a wide set of contexts and countries (Twenge, 2001; Leaper and Ayres, 2007, 357). As politicians are likely to reflect the characteristics of the population from which they are drawn, if women in the UK are now more agentic on average, we should expect these changes to be reflected in the behaviour of politicians too.

Second, changing social roles have affected public perceptions of the validity of traditional gender stereotypes. Women in general are perceived as more agentic now than in the past (Eagly et al., 2020; Sendén et al., 2019), and while attributes associated with men have remained relatively stable, masculine characteristics are increasingly ascribed to women (Diekman and Eagly, 2000). Do changes in attitudes regarding the content and validity of stereotypes mean that voters are less likely to punish counter-stereotypic behaviours? As I reported above, several papers document voters' tendency to punish women candidates for displaying agentic traits. However, I am not aware of any existing empirical literature which tracks the extent to which politicians are punished by voters for contravening stereotypes *over time*. Several more recent studies suggest that voters do not *always* punish women politicians for violating feminine stereotypes (Brooks, 2011; de Geus et al., 2021; Saha and Weeks, 2022), but these papers again only provide evidence from a single point in time.

There is, however, evidence that the public have become less likely to endorse traditional gender stereotypes over time. As women's position in the labour market has improved, support for traditional gender norms and associated stereotypes has eroded both in the UK and further afield (Inglehart and Norris, 2003; Twenge, 1997a; Seguino, 2007). In the UK, voters have come to hold substantially more gender-egalitarian attitudes between the mid-1980s and the present (Taylor and Scott, 2018). Further, between 1990 and 2010, voters in Western Europe, including in the UK, have become significantly less likely to agree with the traditional division of social roles performed by men and women (Shorrocks, 2018). This latter finding is particularly relevant given that it is the association of men and women with particular social roles that is at the heart of theories of gender stereotypes (Eagly and Karau, 2002). It therefore seems likely that as voters have become less willing to endorse gender stereotypes, they also will apply fewer sanctions to politicians who transgress such stereotypes. Consequently, as Mo (2015, 360) argues, "gender attitudes in the electoral process remain consequential, but have grown subtler". To the extent that politicians in the UK are sensitive to the expectations of voters, then, changing voter



attitudes about stereotypes will likely have reduced pressures on women politicians to conform to gender-stereotypic behaviours over time.

Third, the dramatic shifts in the roles that women play in *political* life in recent decades might also reduce the degree to which women politicians conform to traditional gender stereotypes. In the House of Commons, women held just 18% of seats in 1997, but this increased to 32% by 2019 (IPU, 2022). Moreover, women politicians in the UK now occupy more high-powered positions within the legislative hierarchy (Blumenau, 2021c). As women enter politics at a higher rate, role theory predicts that women politicians will come to be seen as possessing more masculine characteristics (Diekman et al., 2005), and the increasing prevalence of women in leadership has been shown to reduce the degree to which communal qualities are ascribed to women (Dasgupta and Asgari, 2004). As Diekman et al. (2005, 212) argue, “women’s increased representation as elected officials and government employees should foster the ascription to women of traditionally masculine qualities.”

Accordingly, in addition to a general tendency for stereotypes of women to become more oriented towards agentic characteristics in recent years, women politicians *specifically* may have become associated with more masculine characteristics over time as they have become more numerous and more powerful in the (historically male) political domain. These shifts in the political sphere might help to further reduce voter sanctions against women politicians who adopt agentic styles, but they are also likely to reduce the degree to which women in parliament *internalise* expectations of feminine behaviour. That is, as women politicians witness more examples of women in politics adopting more agentic and less communal styles, this may weaken the self-imposed standards of femininity that are typically seen as the internal drivers of stereotype-consistent behaviour (Eagly and Wood, 2012). As women politicians become increasingly associated with forms of behaviour normally ascribed to men, the incentives for them to conform to more traditional feminine styles should be expected to decrease.

In the empirical analysis below, I do not attempt to disentangle which of the three mechanisms outlined here, or others, may be responsible for changes in parliamentary behaviour. Rather, I aim to test a central prediction that emerges from the discussion of dynamic gender stereotypes: that MPs will conform less to gender-stereotypic styles in recent years than was true in the past, and that women in particular will be more likely to adopt agentic rather than communal styles over time.

#### *Pressures to conform to institutional behavioural norms*

In this section, I contrast the predictions of my argument with expectations generated by theories of feminist institutionalism. These perspectives hold that, as historically male-dominated institutions,

legislatures are gendered spaces that maintain, favour and recreate traditional masculine behaviours (Hawkesworth, 2003; Krook and Mackay, 2011). While work in this literature does not always articulate clear predictions regarding the dynamics of gendered behaviour over time, implicit in these arguments is the idea that the pressure for women to conform to the prevailing (male) institutional style will be strongest when women are more marginalised in the legislature. When this is the case, as Franceschet (2011, 66) argues, “women may respond by disavowing distinctly feminine (and feminist) concerns, instead favouring the style and substantive issues of the dominant group.” By contrast, as women gain higher levels of representation and more political power, the culture of parliament will change to be more “conducive to women acting in a feminised way” (Childs, 2004b, 187).

The implication of this argument in my setting is that the pressure on women to conform to the dominant “masculine” institutional style will be strongest at the beginning of the study period (in the late 1990s) when women’s representation in the Commons was at lower levels, but that it should weaken over time. In addition to the increasing number of women in parliament and in leadership positions, during the period I study (from 1997 to 2019) the House of Commons also introduced a series of reforms designed to strengthen the position of women within the legislature.<sup>3</sup> Therefore, while the Commons remains majority male, institutionalist perspectives predict that changes in composition and working practices will mean that women will be better able to “perform their tasks as politicians the way they individually prefer” (Dahlerup, 2006, 519). Therefore, while my argument suggests that women are likely to respond to changing gender stereotypes by adopting more *agentic* styles over time, the institutionalism argument suggests that women are likely to adopt increasingly *communal* styles as they become more numerous and powerful in parliament. The empirical strategy I outline below allows me to adjudicate between these contrasting predictions.

### 3.3 Data and methodology

I consider the words of politicians’ speeches as the primary locus of debating style, and I use texts of political speeches delivered in parliament to infer the styles adopted by different speakers. Parliamentary speech is a useful source of information for measuring style as it provides long-running panel data at the individual level. In the UK, MPs are afforded a large degree of autonomy regarding the debates to which they contribute, and party leaders exert no control over who participates, nor

---

<sup>3</sup>For example, this period includes the establishment of the Women and Equalities Committee, the introduction of the Speaker’s Reference Group on Representation and Inclusion, as well as the introduction of initiatives such as proxy voting for MPs on baby leave.

over the content of speeches that MPs deliver.

I study House of Commons debates between May 1997 and March 2019 (Blumenau, 2021a). The study period is motivated by the fact that prior to the 1997 election women accounted for less than 10% of MPs, and so analysis of earlier periods would likely be sensitive to the styles of only few specific women. I collapse the data such that all speeches made by an MP in a debate constitute a single speech-document, making the unit of analysis an individual MP in a debate. I remove all speech-documents shorter than 50 words, as well as contributions by the Speaker of the House, whose speeches are almost entirely procedural. I also exclude any debate that has fewer than five participants.<sup>4</sup> The final sample consists of 14,864 debates, 1,422 MPs (370 women, 1,052 men), and 418,147 MP-debate observations.

#### *Measuring “style” with context-specific dictionaries*

A common approach to measuring latent concepts, such as style, in text data is to assign each text a score based on a predefined dictionary that aims to capture the concept of interest. However, dictionary-based approaches are highly domain-specific, as the words used to capture a concept in one context – say, parliamentary speeches – are likely to be different to those used to express the same concept in another context. I propose an alternative approach that combines standard dictionaries with a locally-trained word-embedding model to construct domain-specific dictionaries that are better able to capture the style types *as they manifest in the context of parliamentary debate*. The key advantage of my approach is that it allows me to account for context-specific patterns of word use. That is, rather than simply using an off-the-shelf dictionary that may be poorly suited to capturing, for instance, aggression in the parliamentary setting, this approach allows me to automatically create a bespoke aggression dictionary which is firmly rooted in the way that vocabulary is used in parliamentary debate. I use this approach to measure six of the styles: aggression, affect, positive emotion, negative emotion, fact, and human narrative.

For each style, I follow three steps to construct the relevant score for each speech. First, I define a “seed” dictionary that represents the concept of interest. For four styles (affect, negative emotion, positive emotion, and fact), I use existing dictionaries and for two styles (aggression and human narrative) I create my own seed dictionaries based on a close reading of a sample of parliamentary texts.<sup>5</sup>

Second, I estimate a set of word-embeddings using the GloVe model described by Pennington,

<sup>4</sup>The model becomes computationally burdensome with very large numbers of debates. Small debates contribute little to the estimates given the random-effect structure of the model described below, and so the results are very unlikely to be sensitive to this decision.

<sup>5</sup>I include a full description of the seed dictionaries in appendix A.1.

Socher and Manning (2014). Word-embedding models rely on the idea that words which are used in similar contexts will have similar meanings, and the embedding model allows us to *learn* the semantic meaning of each word directly from how the word is used by MPs in debate. I train the embedding model on the full set of parliamentary speeches, and the main output of the model is the set of word-embeddings themselves. These are dense vectors that correspond to each unique word in the corpus, the dimensions of which capture the semantic “meanings” of the words. Crucially for my purposes, the distances *between* word-vectors have been shown to effectively capture important semantic similarities between different words (Mikolov et al., 2013).

By calculating the cosine similarity between every word in the corpus and the words in each of the seed dictionaries, I can therefore use this property to define the set of words that, *in the specific context of parliamentary debate*, are used in a semantically similar fashion to the seed words. I label this quantity as  $Sim_w^s$ , where  $w$  indexes words, and  $s$  indexes each style. Words closely related to the average semantic meaning of the seed words for a given dictionary will have a high similarity score (close to 1), and words that are less closely related will have a low similarity score (close to 0). The  $Sim_w^s$  scores therefore define a domain-specific dictionary for a given style type. They describe the degree to which each unique word in the corpus is used similarly to the ways in which the words in the seed dictionary are used, on average. In essence, the embeddings enable the seed dictionaries to automatically expand to incorporate words that are used in a similar manner to the words that they already include. I provide full details of the approach, and an extensive set of validation checks, in appendix A.2.

Third, I use the word-level scores,  $Sim_w^s$ , to score each *sentence* in the corpus on each style according to the words they contain. In particular, the score for a given sentence on a given style is:

$$Score_i^s = \frac{\sum_w^W Sim_w^s N_{wi}}{\sum_w^W N_{wi}} \quad (3.1)$$

where  $Sim_w^s$  is the similarity score defined above, and  $N_{wi}$  is the (weighted) number of times that word  $w$  appears in sentence  $i$ , where the weights are term-frequency inverse-document-frequency weights.<sup>6</sup> When words with high scores for a given style appear frequently in a given sentence, the sentence will be scored as highly relevant to the style. The score for each *document* is then the weighted average of the relevant sentence-level scores, where the weights are equal to the number of words in each sentence.

<sup>6</sup>TF-IDF weighting is used to down-weight very common words, and up-weight relatively rare words.

The approach I outline here is similar to that developed in [Rice and Zorn \(2021\)](#), who also use a word-embedding model to create context-specific dictionaries. I build on this work by addressing the question of whether word-embedding dictionary construction “[yields] valid dictionaries for widely-varying types of specialised vocabularies” ([Rice and Zorn, 2021](#), 34). I extend the idea of word-embedding based dictionaries to a new setting – the UK House of Commons – and to six new specialised vocabularies. As the validation exercises in appendix [A.2](#) show, there is strong evidence that this approach significantly outperforms standard dictionary approaches across the set of styles I study.

#### *Measuring “complexity” and “repetition”*

For the final two styles – complexity and repetition – dictionary approaches (domain-specific or otherwise) are unsuitable, as these styles are not detectable from the occurrence of specific words. Instead, I adopt two different metrics to capture these concepts. For *complexity*, I use the Flesch-Kincaid Readability Score ([Kincaid et al., 1975](#)). The intuition behind this measure is that documents that have fewer words per sentences, and fewer syllables per word, are easier to understand (more “readable”). I rescale the original formulation of the score such that higher numbers indicate higher levels of complexity. While [Benoit, Munger and Spirling \(2019, 501\)](#) show that domain-specific measures of textual complexity have some performance gains over the Flesch-Kincaid score, they also demonstrate that this metric correlates highly with more sophisticated measures. I opt for the simpler metric here because the validation demonstrates that this measure performs well in comparisons with human judgments in this setting.

I consider MPs to be *repetitive* when they use the same language repeatedly during a debate. To measure repetition, I use a lossless text-compression algorithm introduced by [Ziv and Lempel \(1977\)](#), which underpins a variety of common computer applications. Compression algorithms work by finding repeated sequences of text and using those patterns to reduce the overall size of the input document. The efficiency of the compression of a text is directly related to the number and length of the repeated sections in that text. I apply the compression algorithm to every document in the corpus, measure the degree of compression, and treat that quantity as the measure of repetition for each MP in each debate. Simply put, the more compression that a speech receives, the more repetitive I deem it to be.<sup>7</sup>

Finally, to put the eight style measures on comparable scales, I normalise each measure across

---

<sup>7</sup>An alternative measure of repetition for a given text,  $j$ , might be  $\frac{\# \text{Words}_j}{\# \text{Unique Words}_j}$ , which captures the intuition that texts with a smaller fraction of unique words are likely to be more repetitive. Although it correlates highly with my measure ( $\rho = 0.71$ ), this metric is likely to underestimate the degree of repetitiveness in instances where long sequences of words are repeated, but where those sequences are themselves constituted of many unique words.

documents to have mean zero and standard deviation one. This means that average differences for each style between men and women can be interpreted in standard deviations of the outcome variable.

### *Validation*

As with all quantitative text analysis approaches, careful validation of the measures is essential (Grimmer and Stewart, 2013). In the appendix in section A.2, I provide the results from two face validation checks. I here provide results from a human validation task which assesses whether the text-based measures of style mirror human judgements of the same concepts. I wrote a web app which presented two research assistants with pairs of sentences (sampled from all sentences in the corpus). Coders were asked to complete two tasks. First, a *style-comparison* task required them to select which of the two sentences was more typical of a particular style. Second, a *style-intensity* task required them to rate the degree to which each sentence was representative of the selected style on a 5-point scale.

Figure 3.1 gives an example of the prompt seen by the coders. In addition to the sentences themselves, I presented coders with minimal definitions of the speech-styles of interest to ensure that the human coding related to the style dimensions identified in the literature review. Each coder completed 70 comparisons per style, on average, meaning that I have on average 140 individual sentence-ratings per style. I use the distribution of responses to these tasks and compare them to the distribution of text-based style measures described above for the same sentences as seen by the coders.<sup>8</sup> I summarise the results in table 3.2. The “intensity task” column presents the correlation between the sentence-level style measures (equation 3.1) and the coders’ ratings of the same styles. For the “comparison task” column, I calculate the difference in the sentence-level scores for each pair of sentences, and correlate that with the choices made by my coders from the comparison task.

---

<sup>8</sup>To assess inter-coder reliability, my research assistants both coded an additional common set of 20 comparisons per style. Coders agreed on which of the two sentences was more representative of a given style in 75% of comparisons. The correlation for the “intensity” scores for all sentences across coders was 0.8.

# Style Validation

Introduction

Validation

Progress

## Fact

Your task is to select the sentence which you believe uses more **factual** language, which might include the use of numbers, statistics, numerical quantifiers, figures and empirical evidence.

### Sentence one

Lower than expected unemployment is already saving around £10 billion over the next five years on benefit spending alone, compared with Budget plans.

### Sentence two

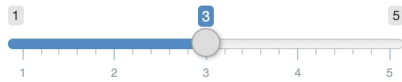
Credit unions and money advice centres also deal with several thousand similar cases each year.

Which of these sentences uses more **fact-based** language?

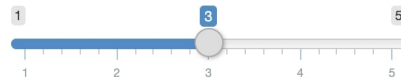
- Sentence one.  
 Sentence two.  
 About the same.

On a scale where 1 is not at all representative of **fact-based** language and 5 is very representative of **fact-based** language, where would you place...

...sentence one?



...sentence two?



Next

Figure 3.1: Human validation task prompt

Table 3.2: Correlation between text-based measures and human judgements

Style type	Comparison task	Intensity task
Human Narrative	0.67 (0.5)	0.7 (0.45)
Affect	0.62 (0.46)	0.61 (0.48)
Positive Emotion	0.7 (0.38)	0.71 (0.34)
Negative Emotion	0.75 (0.47)	0.75 (0.45)
Fact	0.77 (0.71)	0.81 (0.74)
Aggression	0.66 (0.32)	0.72 (0.22)
Complexity	0.83	0.85
Repetition	0.8	0.82

Overall, the results are very encouraging. Across all styles, the correlation between the text-based scores and the human validation is always positive and is never lower than 0.61 for either task. These

results suggest that there is a clear correspondence between the measures of style implied by my text-analysis approach, and human judgements of those concepts in the same set of texts.<sup>9</sup>

Moreover, I can compare the measures with standard dictionary-based measurement approaches. For all styles except for repetition and complexity, I compare the word-embedding approach to an approach that measures style using the proportion of words in each sentence that appears in a pre-defined dictionary. This measurement strategy is more typical of existing applications of dictionaries in political science, and forms the basis of the analysis in several previous studies on gender and political style (e.g., Gleason, 2020; Jones, 2016; Yu, 2013). To maximise comparability, the dictionaries I use for this analysis are the same as the seed dictionaries I use to construct the word-embedding scores. This means that, for each sentence in the corpus, I have a measure of style based on the word-embedding method, and a measure of style based on counting the fraction of words in the sentence that fall into the relevant style's seed dictionary.

The results are given in table 3.2. The numbers in parentheses show the correlation between the standard dictionary measure of style described above, and human judgements provided by my coders. My word-embedding approach clearly outperforms standard dictionary approaches in approximating human judgement. For instance, for positive emotion, standard dictionary measures correlate at 0.38 and 0.34 with human codings for the two tasks, compared to 0.7 and 0.71 for the word-embedding approach. Despite the relatively small sample sizes, the magnitude of the difference in predictive power means that – in all cases except for “fact” – the correlation between the word-embedding measures and human codings is significantly higher than the equivalent correlation for standard dictionary measures.<sup>10</sup> Overall, this exercise provides strong evidence that I can reliably detect the styles of interest in parliamentary speech and outperform the standard measures used in previous studies on gender and political style.

### 3.4 Modelling political style

My goal is to assess the degree to which style use varies by MP gender, and whether such differences change over time. To investigate these patterns, I adopt a Bayesian dynamic hierarchical model that

<sup>9</sup>As repetitiveness is a quantity that manifests more clearly *across* rather than *within* sentences, the sentence-based human validation is somewhat less well suited to evaluating this concept. Nevertheless, the sentences that the measure marks as most repetitive do clearly demonstrate high levels of repetitiveness, and, as table 3.2 indicates, even though detecting repetitiveness at the sentence-level might represent a hard task, I recover a clear correspondence between my measures and human judgements of the same concept.

<sup>10</sup>I determine this difference by using a bootstrap procedure, in which I sample from my set of sentences 2000 times with replacement and calculate the correlation between the word-embedding measures and human codings, and between the dictionary measures and human codings, on each iteration. I can easily reject the null hypothesis of no difference in these correlations for all styles except for the “fact” dimension.



allows me to account for a wide variety of both individual- and topic-level confounders (described below), while also flexibly estimating changing gender dynamics in style over time.

For each speech  $i$ , I have a continuous measurement of style  $s$ , which I denote as  $y_i^s$ . For speech  $i$ , by MP  $j$ , in debate  $d$ , and time period  $t$ , I model the data as a function of individual- and debate-level parameters:

$$y_{i(jdt)}^s \sim N(\alpha_{j,t} + \delta_d, \sigma_y) \quad (3.2)$$

where  $\alpha_{j,t}$  is an MP-specific random effect which captures average differences in MP style use. The  $t$  subscript indicates that I fit one intercept for each MP in each time period that they appear in the data, thus allowing me to capture average style use at different points in time. I use parliamentary sessions as the unit of time, of which there were 20 between 1997 and 2019. I observe speeches from 635 MPs on average in each session, and each MP appears in 9 sessions on average. The  $\delta_d$  parameters are random effects which capture average differences in style use in different debates.

My primary interest is in describing variation in the  $\alpha_{j,t}$  parameters. I model these random-effects at the second level of the model as a function of MP gender, while allowing the relationship between gender and style-use to vary over time:

$$\alpha_{j,t} \sim N(\mu_{0,t} + \mu_{1,t} \text{Woman}_j, \sigma_\alpha) \quad (3.3)$$

Here,  $\mu_{0,t}$  represents the average use of a style among men in time period  $t$ , and  $\mu_{1,t}$  describes the average difference in style use for women relative to men, again in time period  $t$ . The standard deviation  $\sigma_\alpha$ , describes how much, on average, the MP-session intercepts vary around the mean style use for MPs of each gender.<sup>11</sup> Gender differences in one parliamentary session are not independent of those in previous sessions, and in order to reflect a more realistic evolution of these differences, I model the  $\mu_{0,t}$  and  $\mu_{1,t}$  parameters as a first-order random-walk process:

$$\begin{aligned} \mu_{0,t} &\sim N(\mu_{0,t-1}, \sigma_{\mu_0}) \\ \mu_{1,t} &\sim N(\mu_{1,t-1}, \sigma_{\mu_1}) \end{aligned} \quad (3.4)$$

This specification assumes that the average use of a style by women and men will be similar in  $t$  and  $t+1$ , and that changes over time will therefore occur gradually. This encourages smooth coefficient

<sup>11</sup>I use a common variance parameter for all time periods.

changes over time, but still allows for large deviations from one period to the next if the information from the data is sufficiently strong.

The  $\mu_{0,t}$  and  $\mu_{1,t}$  parameters are my main quantities of interest.  $\mu_{1,t}$  captures the difference in average style use between genders in each time period, and my review of the theoretical literature implies general expectations for the *sign* of  $\mu_{1,t}$  for each style (see table 3.1). Consistent with my theoretical discussion of how the incentives for conforming with gender stereotypes have changed in recent years, I also expect the *magnitude* of  $\mu_{1,t}$  for each style to decrease over time, and for those changes to be driven mostly by changes to the average behaviour of women MPs (which, for each year  $t$ , is captured by  $\mu_{0,t} + \mu_{1,t}$ ). I report both quantities below.

My model allows me to account for individual-level confounders by including a set of MP-specific covariates into the model. To do so, in some specifications I replace equation 3.3 with:

$$\alpha_{j,t} \sim N(\mu_{0,t} + \mu_{1,t} \text{Woman}_j + \sum_{k=1}^k \lambda_k X_{j,t}^k, \sigma_\alpha) \quad (3.5)$$

where  $X_{j,t}$  is a vector of individual-level covariates which can vary by session.<sup>12</sup>

I include several such controls. First, MPs in leadership positions may use systematically different styles than backbench MPs, and women have come to occupy a greater share of legislative leadership roles over the study period (Blumenau, 2021c). I therefore control for whether the MP held a frontbench position for either the government or opposition in each session, and whether they were a committee chair. Second, if MPs from different parties use styles at different rates, then any change I observe in gendered use of styles might be confounded by the fact that proportionally more Conservative Party women MPs have been elected to parliament in recent years. I therefore also include a set of party dummies. Third, opposition MPs use significantly more negative language than government MPs (Proksch et al., 2019) and, because the Labour Party has proportionally more women than other parties, any increase in women’s use of more agentic styles might be attributable to Labour’s move into opposition in 2010. To address this possibility, I control for whether an MP is a member of a governing or an opposition party in each time period. Fourth, I also add controls for MPs’ occupational background

<sup>12</sup>Debate intercepts are drawn from a mean-zero normal distribution, with estimated variance:

$$\delta_d \sim N(0, \sigma_\delta) \quad (3.6)$$

Our model is completed by normal prior distributions over the  $\lambda$  parameters:

$$\lambda^k \sim N(0, 2) \quad (3.7)$$

and half-normal prior distributions over the variance-parameters:

$$\sigma_\alpha, \sigma_{\mu_0}, \sigma_{\mu_1}, \sigma_\delta, \sigma_y \sim N(0, 2) \quad (3.8)$$

and educational attainment. The professional and educational backgrounds of MPs have changed over time (Lamprinakou et al., 2017), and it is plausible that these characteristics will be associated with both speechmaking styles and gender. Finally, MPs' local electoral environment might affect language use. For instance, MPs in more competitive seats might be more likely to use human narrative to emphasise constituents' concerns. If there have been changes in the relative competitiveness of seats won by men and women over time, this could confound the differences I observe in gendered-language use. I therefore control for the percentage point margin of victory of the MP in the previous election.

I am also able to use the model to account for confounding that relates to differential usage of styles across topics. As I later show in chapter 4, men and women systematically participate in debates devoted to different topics, and debate topic may correlate with style in ways which work to confound my inferences. For instance, if women participate more in debates on education which contain language related to human narrative, while men participate more in debates on the economy which include more factual language, then gender differences in topic usage will confound gender differences in style. However, the debate-level intercepts,  $\delta_{d,t}$ , mean that it is only *within-debate* variation in style use that informs the estimates of my central quantities of interest. In other words,  $\mu_{1,t}$  will capture only the degree to which men and women use different styles when speaking about the same substantive topic.

I estimate the model separately for each style in Stan (Carpenter et al., 2017), where I use three chains of 500 iterations, after 250 iterations of burn-in.

### 3.5 Results

Figure 3.2 shows the values of  $\mu_{1,t}$  – the average difference between men and women for each style type, in each parliamentary session. Positive values indicate that a style is used more by women, and negative values indicate higher use of the style by men. The green shading indicates the expected direction of the gender effects based on previous literature (see table 3.1).

For five of the debating styles I study, I find that – in the early years of the sample – men and women's speechmaking behaviour broadly conforms to stereotypes. Women MPs are more likely to draw on examples that emphasise human narrative, and to use positive and emotional language than men. Similarly, men use more aggressive and complex language, at least before 2010. Interestingly, for three of the styles – fact, repetition, and negative emotion – I find that debating style in the Commons does not clearly conform to the expectations of the existing literature. For all three of these agentic

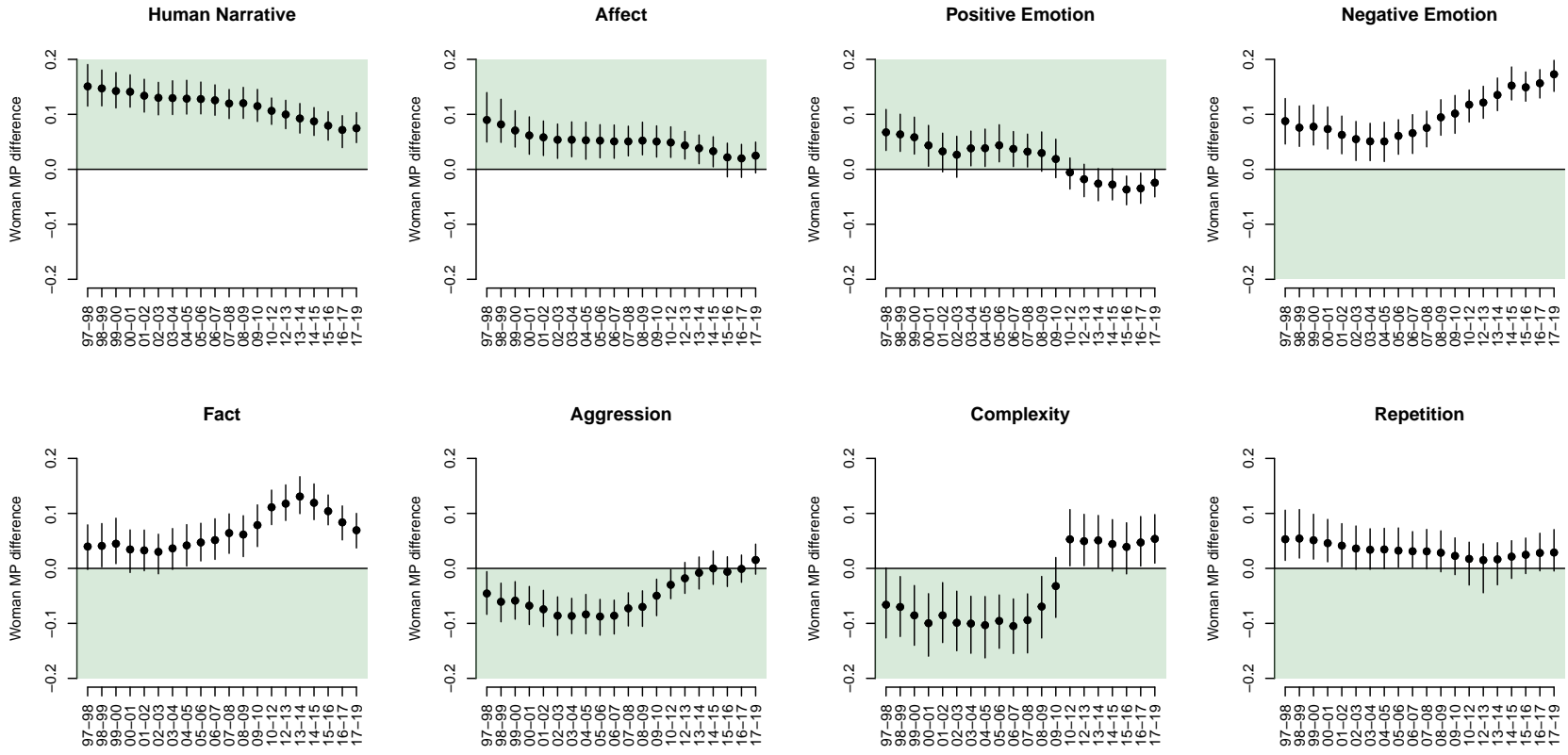


Figure 3.2: Gender differences in style over time

styles, for much of the period I study, women MPs are *more* likely than men to express these styles.

However, and in some sense more importantly, figure 3.2 also reveals that there is significant variation in the size of these gender differences over time. Women are more likely to use communal style types – affect, positive emotion, and human narrative – in the early period in my data, but gender differences become smaller over time. For positive emotion and affect, there is no consistent significant difference between men and women by the latest years in my data. Similarly, while men use significantly more agentic styles – particularly aggressive and complex language – than women before 2010, this difference has also disappeared in recent years. These changes are non-trivial: for those styles where I see a convergence between men and women, the proportion of MP-level variation in style use explained by gender decreases by between 30% and 60%, depending on style, when comparing the periods before and after 2007.<sup>13</sup>

Further, for other styles I observe increasing gender differences over time, but the direction of these shifts also suggest that women are becoming increasingly agentic relative to men. For instance, though there are negligible gender differences in the earlier period, from 2007 onward, women use significantly more factual language. Similarly, although women use negative emotion in their speeches at higher rates throughout the time period, this gender difference has grown substantially larger over time. Between 1997 and 2007, gender explained just 0.6% and 2% of MP-level variation in factual language and negative emotion, respectively, but this increased to approximately 4% and 9% after 2007. Accordingly, even for these agentic styles which women adopt more than men throughout the study period, it remains the case that women become *more* likely to deploy this type of language in recent years than in the past. The only style for which I document relatively stable gender differences is repetition. While women appear to become somewhat less repetitious relative to men over time, the trend for this style is less pronounced.

Taken together, these findings are consistent with my argument that the pressures for women to conform to stereotypes have declined over time. In general, relative to men, women demonstrate less communal (human narrative, affect, and positive emotion) and more agentic (negative emotion, aggression, fact, and complexity) styles in recent years than they did in the past.

<sup>13</sup>To calculate these quantities, I follow Gelman and Pardoe (2006) and describe the proportion of individual variation in style use explained by MP gender in each parliamentary session using an  $R^2$ -style metric:

$$R_{\alpha,t}^2 = 1 - \frac{E(V_{j=1}^J \hat{\epsilon}_{j,t})}{E(V_{j=1}^J \hat{\alpha}_{j,t})}$$

where

$$\hat{\epsilon}_{j,t} = \hat{\alpha}_{j,t} - \hat{\mu}_{0,t} + \hat{\mu}_{1,t} \text{Woman}_j$$

Are these patterns the result of changes in the behaviour of men and women MPs? My argument implied that these changes were likely to be rooted in the behaviour of women as they respond to the changing content and power of gender stereotypes. Figure 3.3, which depicts changes in style use separately for women and men, shows that across almost all the styles I study, the largest year-to-year shifts in speechmaking behaviour do indeed occur among women. The figure shows that women have used each of the communal style types – human narrative, affect, and positive emotion – to a decreasing extent over time. Similarly, for the agentic styles of negative emotion, fact, and aggression, while men’s behaviour has remained relatively stable, women’s use of these styles has increased over time. While both men and women have adopted more complex language over time, the increase has been somewhat larger for women than for men.

#### *Threats to inference*

I have argued that politicians’ conformity to traditional stereotypes has diminished over time, but there are alternative explanations that could account for the behavioural patterns I document.

First, as I outlined above, the literature on feminist institutionalism suggests that women will face pressures to conform to institutionally-dominant, masculine behaviours that are favoured and recreated by the culture of the House. While institutional pressures of this sort are surely a feature of life in the contemporary House of Commons, for this perspective to explain the results, these pressures would need to have *strengthened* in recent years. However, during this period, women’s presence increased in the Commons, women MPs came to hold more senior positions, and institutional reforms designed to strengthen women’s institutional position were introduced. Consequently, if women face stronger incentives to conform to masculine styles when an institution is more male-dominated, then we should observe women becoming *less* agentic over time. My results show the opposite pattern, suggesting that institutionalist accounts are unlikely to explain the over-time dynamics in speechmaking that I document.

Second, figures A.3 and A.4 in the appendix report results from the model described in equation 3.5 in which I control for a host of MP-level covariates. If the changes I observe over time are driven by factors such as party, opposition status, and so on, I would expect to see large differences between these two sets of results. Although there is some attenuation of the over-time changes for aggression and complexity when controlling for covariates, I nonetheless still observe stereotype-consistent differences at the beginning of the sample period and clear evidence of women using more agentic styles over time.

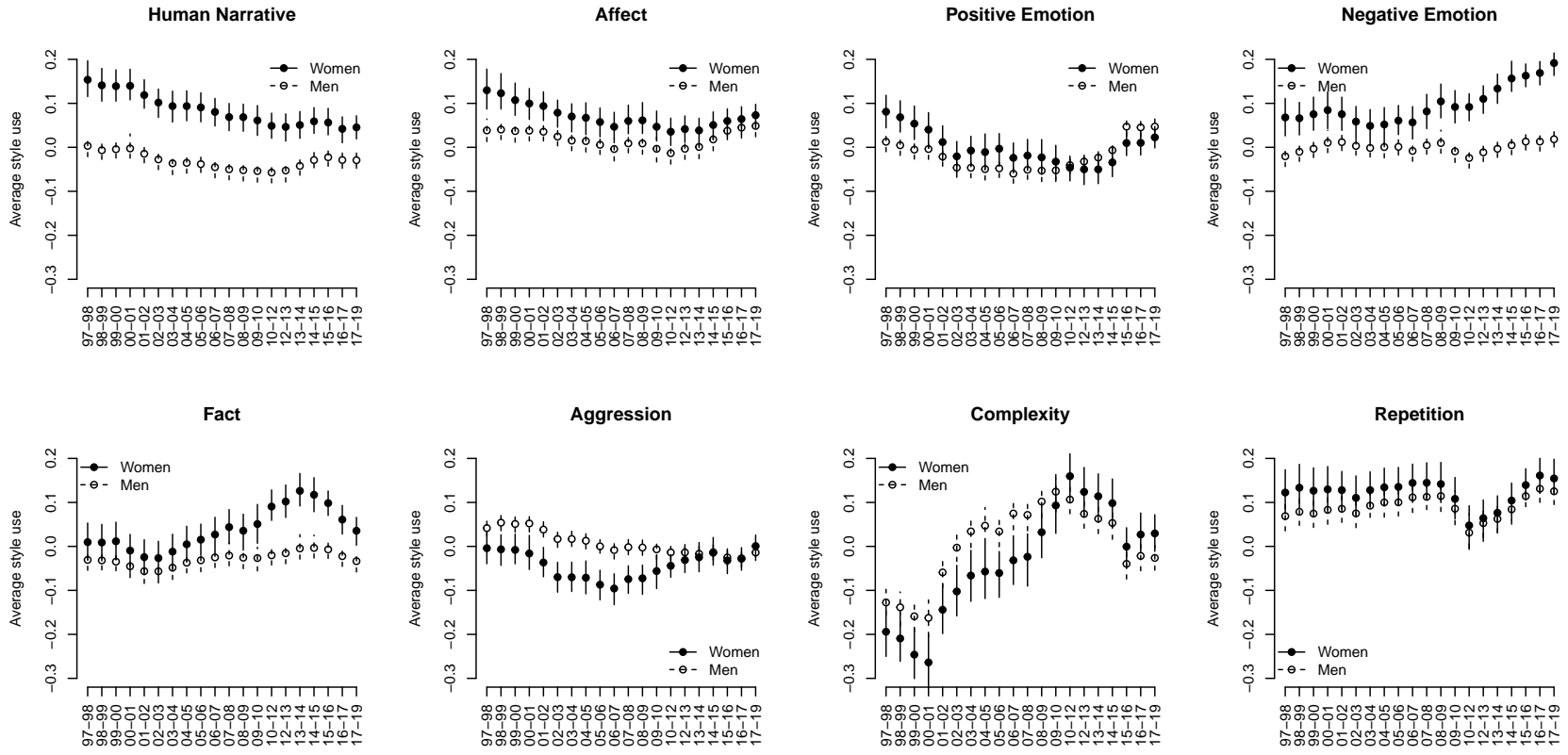


Figure 3.3: Average style use for men and women over time

Third, the aggregate patterns I document could reflect changes to the parliamentary agenda, rather than changes in gendered behaviour. If there are certain topics on which women are more likely to demonstrate agentic styles, and these topics feature more prominently on the parliamentary agenda in the later period, then my results might be explained by changing topical prevalence over time. For instance, I observe a large converge between men and women's use of human narrative over time. Women are significantly more likely to use human narrative in their parliamentary speeches at the beginning of the time period than they are at the end. If, however, women are more likely to use human narrative than men in certain *topics*, and those topics become less prevalent over time, then the convergence I document might in fact be attributable to changes to the parliamentary agenda. For changes in topic prevalence to be responsible for convergence, it would have to be the case that the topics on which I observe women using more human narrative than men are becoming *less* prevalent, or that the topics on which I see women using less human narrative than men are become *more* prevalent over time. For example, perhaps women use more human narrative than men when discussing education policy, and education policy is more frequently discussed in the early period in the data than the later period in the data. If this were true, then my results might be subject to topical confounding, as changes in topical prevalence over time would account for the aggregate changes I report above. In appendix A.6, I use statistical topic-models to measure the differences between men and women's style use across a wide variety of topics, and then evaluate whether topics that are marked by large stylistic differences become more or less prevalent over time relative to topics marked by smaller differences. I find scant evidence of such topical confounding.

Finally, I show that, on average, women MPs deliver speeches that are less likely to be marked by communal styles and more by agentic styles over time. One potential alternative explanation for these results is that men and women MPs who employ different speaking styles might have become differentially likely to *participate* in parliamentary debate over time. We might imagine, for instance, that women MPs who tend to deliver highly agentic speeches gave more speeches in parliament over the course of the study period, and that women who tend to deliver highly communal speeches participated less in debate over time. If that were the case, differential participation might drive the changing gendered speechmaking dynamics that I document above, rather than within-MP changes. In appendix A.7, I assess whether women with more agentic styles participate more, and women with more communal styles participate less, in parliamentary debate over time. I show that differential participation does not explain my results: MPs' styles largely fail to predict debate participation throughout the study period. In addition, I also investigate whether the changes I document are due to changes in styles



of men and women throughout their careers (“within-MP” effects), or because the men and women entering parliament over time are systematically different from those leaving (“replacement” effects). While there is some evidence that replacement is more important for explaining the changes in agentic styles and within-MP change is somewhat more important for explaining change in communal styles, overall, I find that neither replacement nor within-MP change can alone explain the patterns that I document above.

### 3.6 Conclusion

My central substantive contribution is to document the fact that gender stereotypes are worse descriptors for actual political behaviour in the UK now than was true in the past. In particular, in recent years, women in the House of Commons demonstrate less communal and more agentic styles, and the gender gap on most dimensions of style that I examine has decreased. I see these results as an important corrective to the scholarly literature on gender differences in legislative behaviour, which typically emphasises that men and women politicians argue in ways that are broadly consistent with stereotypes. Though this may be true in some settings, gender stereotypes of communication styles have become significantly less predictive of the reality of contemporary British political debate.

These findings do not, however, imply that gender-stereotypes play no role in UK politics. I show that recent parliamentary behaviour is poorly described by traditional stereotypes, but I do not provide empirical evidence regarding the mechanisms that led to these changes. For instance, previous work shows that the public are less likely to endorse traditional stereotypes now than in the past, but I lack data to assess whether there has been a concomitant decline in the sanctions that voters apply for gender-role-inconsistent behaviour. Anecdotally, there continue to be examples of British women politicians being criticised for stereotype-incongruent behaviour<sup>14</sup> and stereotypes may well continue to condition voter responses. However, in chapter 5, I present evidence that is consistent with the results I present here: that women politicians in the contemporary UK are *not* punished for role-inconsistent behaviour. Therefore, this type of sanctioning is perhaps likely to have declined because of general changes in voter attitudes regarding stereotypes, future work should nonetheless focus on collecting over-time survey data on voters’ attitudes towards non-stereotypical behaviour by politicians.

One optimistic view of the findings, however, is that there may be a virtuous circle in which women

---

<sup>14</sup>See “The Making of the Maybot”, *Spectator*, 2nd November 2017.

politicians diverge from stereotypical behaviours, and that this in turn changes perceptions of appropriate feminine behaviour, which thereby reduces the pressures women are under to conform to such stereotypes. Changes in the typical behaviours of men and women politicians are likely to translate only slowly into revised public expectations of the standards against which men and women MPs are judged. However, to the degree that the behavioural shifts that I document are noticed and internalised by the public, they might also help to reduce the social penalties applied to women politicians who display more agentic styles.

My findings also have implications for wider debates about political culture in the UK. There is a strand of popular commentary that implies some of the more unattractive features of Westminster's adversarial culture would be ameliorated if only more women were to be elected to public office. My results suggest, however, that simply increasing women's numbers in parliament is unlikely to make UK politics gentler or more deliberative. The pursuit of a "better" politics requires more than vaguely hoping, on the basis of a dogged adherence to outdated gender stereotypes, that the election of women will fundamentally change the ways that our representatives communicate.

Methodologically, this chapter addresses a well-known problem for quantitative text analysis based on dictionaries: the words that demonstrate a given concept in one context may be poorly suited to detecting the use of that concept in another context. I used a word-embedding model to capture how different political styles manifest in the specific setting of parliamentary debate. Results from the validation (in appendix A.2) show that this approach significantly outperforms existing methods, a finding I believe justifies adoption elsewhere. Indeed, in chapter 4, I show that this method is also accurately able to detect the topical focus of MPs' speeches. This strategy is likely to be useful whenever researchers are interested in measuring a latent concept from a large corpus of texts, but where the domain of interest differs from the domain in which existing dictionaries were developed. This describes a large fraction of applications of dictionary methods, and so the approach has the potential to be applied widely elsewhere.

Finally, I focus only on style as expressed in legislative debates. Style may, of course, manifest in other forms of legislative behaviour, or, indeed, other forms of political speech. While I show that gender gaps in legislative speech have declined, it remains possible that gender stereotypes may still be powerful in other arenas, such as in campaign communication where politicians may be particularly sensitive to voter penalties. I hope that my findings motivate other scholars to explore how gender-based differences in political communication have evolved over time in other contexts.

## Chapter 4

# Earning Their Stripes? Gender, Political Experience, and Policy Prioritisation in Parliamentary Debate

### Abstract

How does political experience affect the issues that men and women raise in parliament? Gender stereotypical expectations dictate that women may focus more on traditionally “feminine” policy areas and amplify women’s experiences in the policy-making process. While past work has presented findings broadly consistent with these expectations, it has neglected to assess how women’s policy priorities change with increased political experience. Focusing on the UK, I argue that gender differences in the extent to which politicians raise issues traditionally associated with women will be most pronounced among junior politicians who lack experience in the eyes of the public and their colleagues. However, this will decline with increased political experience. To test my argument, I study parliamentary debates between 1997 and 2019 and use novel quantitative text approaches to measure the issues politicians raise. I find that, among junior politicians, women talk significantly more about “feminine” issues, however this gender gap decreases markedly as seniority increases. Women therefore change the policies they focus on with increased experience, however, I also show that they continue to raise women’s perspectives in this wider set of political issues. These findings have important implications for the representation of women’s voices and perspectives within the policy-making process.

## 4.1 Introduction

How does political experience affect the issues that men and women raise in parliament? Of central concern is the question of whether the demographic characteristics of politicians determine how they behave in office and, specifically, whether they will dedicate greater efforts towards championing political issues that are related to these characteristics and raising the perspectives of their demographic group (Pitkin, 1967). With respect to gender, scholars have identified two primary ways through which men and women differ in their approach to policy prioritisation. The first is that they dedicate more attention to raising substantive issue areas that are related to the stereotypes that arise from traditional gender roles, where women may orientate towards policies that focus on caring for others, such as welfare, health, or education, and men to “tough” and “hard” issue areas, such as the economy, crime, or defence (Eagly, 2013; Schneider and Bos, 2019). The second is that women politicians do more to amplify the concerns, experiences, and perspectives of women in issues across the policy process more broadly (Schwindt-Bayer, 2006). These questions have been addressed empirically by assessing whether women politicians dedicate legislative speeches to (Clayton, Josefsson and Wang, 2017; Dietrich, Hayes and O’Brien, 2019), participate in debates on (Bäck and Debus, 2019; Catalano, 2009), or support legislation on (Piscopo, 2011; Swers, 2002) traditional “feminine” social policy areas and whether they raise the experiences of women more broadly.

This rich literature suggests that, although context may mitigate the link (Childs and Krook, 2009; Hawkesworth, 2003), women politicians are more likely to raise issues traditionally associated with women and amplify the voices of women in the policy process more broadly, and that this is observable in various aspects of their legislative behaviour. However, existing work has not considered how these dynamics change over the course of politicians’ careers. As women have historically been underrepresented numerically in legislatures and have, on average, less experience in office than men, it has been challenging to identify the implications that increased political experience may have for both the behaviour of women legislators in office and for the representation of voters’ interests in the policy-making process. In this chapter, I build on literature on both the dynamics of legislative careers (Bailer and Ohmura, 2018; Benedetto and Hix, 2007; Miquel and Snyder, 2006; Rush and Giddings, 2011) and political experience (Eggers, Vivyan and Wagner, 2017; Fulton, 2012; Lazarus, Steigerwalt and Clark, 2022; Sevi, 2022), in order to understand the effect of politicians’ increased experience on their representational behaviour. In doing so, I shed light on the importance of increasing the presence of women not only numerically in parliaments, but also ensuring they reach sufficiently senior positions.

I argue that gender differences in the extent to which women politicians raise issues traditionally associated with women will be most pronounced among junior politicians, however, this difference will decrease with increased parliamentary prestige for several reasons. First, politicians are at their most inexperienced at the beginning of their careers and will therefore gravitate towards issues where they have the most prior expertise (Bailer and Ohmura, 2018). This experience will be determined by a variety of factors such as prior career, socialisation, and life experiences (Rush and Giddings, 2011), however, for women, this may also be shaped by their shared identity with women in the population (Wängnerud, 2009). Second, when new to these roles, politicians also need to gain credibility and competence in the eyes of voters and their fellow politicians and are therefore incentivised to prioritise issues where they will be accredited the most competence in by others. Crucially, as women have traditionally been perceived as better suited to, and to have greater expertise in, “feminine” areas such as childcare and family issues, education, welfare, and health (Herrnson, Lay and Stokes, 2003), they will place particular emphasis on raising issues that they are likely to be ascribed more expertise in by others. However, as politicians spend more time in office, learn to navigate their roles, and establish themselves in a wider variety of policy areas (Miquel and Snyder, 2006), the selection of issues they raise will diversify as they gain both greater experience in a wider set of issues *and* more credibility in this work from voters and their fellow politicians. I argue that, as a function of their increased experience, the issues that women raise will be determined less by their gender alone, but by a wider variety of factors too.

To evaluate this argument, I focus on the sets of issues that MPs raise when participating in UK House of Commons debates. Work to date has tended to examine whether there are gender gaps in the extent to which women politicians raise women’s perspectives in political debate (Pearson and Dancey, 2011b), or whether they participate in traditional “women’s issues” such as healthcare, education, or welfare (Bäck and Debus, 2019). I contribute to this work by studying both a range of political issues that prior work has identified as stereotypically “feminine” and “masculine” (Krook and O’Brien, 2012) but also a wider set of issues that are central to UK politics (Bevan and Jennings, 2019). In doing so, I follow recommendations from previous literature not to entirely determine *a priori* which sets of issues should be considered to be “women’s interests” (Celis et al., 2014; Yildirim, 2021). The issues I evaluate are agriculture, children and family, crime and policing, defence, education, environment, finance and economy, healthcare, trade, transport, and welfare. To measure the participation of men and women MPs on these issues, I develop new measures by combining traditional text-as-data dictionary approaches with a word-embedding model to construct domain-specific dictionaries that aim

to capture issues as they manifest *in the specific context of UK parliamentary debate*. Further, I also measure the degree to which MPs make direct references to women in their speeches, which is one way of identifying the relative propensity for men and women MPs to raise the perspectives of women in parliamentary debate. I apply these measures to 437,503 speeches delivered by Members of Parliament in the UK House of Commons between 1997 and 2019.

I report four main findings. First, at the aggregate level, the topics that men and women participate in are broadly consistent with the traditional divisions highlighted in prior literature. While for some issue areas, such as the economy or the environment, there are no gender differences, women do talk more about children and family, health, education, and welfare than men, even when accounting for likely sources of confounding such as party and ministerial or committee positions. Second, these aggregate gender differences mask significant variation over the course of politicians' careers. Among junior politicians, women talk significantly more about traditionally "feminine" policy areas, such as children and family, health, education, and welfare, however, the gender gap reduces markedly among senior politicians. The magnitude of these changes is large: comparing new MPs with MPs with 9 years of experience, the median experience level in my study period, the gender difference in raising issues related to education decreases by 47%. Similarly, there is a 32% reduction in the size of the gender gap for health, 19% for children and family, and 28% for welfare. Further, this pattern is concentrated only among these policy areas, and not among the wider sets of issues politicians raise.

Third, I find that women politicians amplify the voices of women to a greater extent than men, and that this difference does not change across the career cycle. Finally, I find that this difference holds across debates on all policy areas. Indeed, the largest differences I report are concentrated among traditionally "masculine" policy areas, such as crime and policing or defence. This suggests that when participating in debates on crime and policing, women politicians raise the experiences of women to a greater extent than men, and potentially highlight gendered aspects of these issue areas, such as the failure to adequately police violence against women. Further, I find that this, too, does not shift with political experience. Women therefore diversify the sets of policy areas they focus on the longer they are in parliament, however they also continue to raise women's perspectives and experiences in this wider set of political issues. These findings have important normative implications, as they suggest that it is vital not only to incorporate women numerically into the policy-making process, but to ensure that they serve sufficiently long parliamentary careers to best substantively represent the experiences and perspectives of women in debates across the policy spectrum.

This chapter contributes to the existing literature in several ways. First, while evidence from politi-

cians' testimonies suggest men and women MPs differ in the issues they raise in parliamentary debates (Childs, 2004b), with few exceptions (Bird, 2005; Catalano, 2009, described in further detail below), there is currently only limited systematic evidence that this is the case in the UK. I examine gender divisions in the issues men and women raise over the past 22 years and show that men and women *overall* raise different issues in parliamentary debates, and that these differences broadly coincide with the expectations of prior literature. Second, I show that these gender differences on the stereotypically "feminine" policy areas decrease with increased parliamentary experience. In doing so, I contribute to a growing body of literature that has sought to understand how incentives shape politicians' representational efforts (Bailer et al., 2022; Höhmann and Nugent, 2021; Weeks et al., 2022), by arguing that incentives shift in politicians over time because of the experience and credibility they gain throughout their careers. Third, I show that women politicians do more to elevate women's voices across debates on a wider variety of political issues, and that this is unaffected by their parliamentary career cycles. The results I present have positive implications for both the experiences of women politicians in office, but also for the substantive representation of women's interests in politics more broadly. As women politicians become more senior, they champion a wide variety of issues beyond those traditionally associated with women, and continue to raise women's experiences and perspectives in these issues, thus ensuring that the interests of women voters are raised in debates across the policy process. Taken together, these findings have important implications for the representation of women's voices and perspectives in politics.

#### **4.2 Descriptive and substantive representation: Evidence and explanations**

Descriptive representation refers to the idea that a legislature is comprised of politicians that reflect the characteristics of the wider population. Substantive representation refers to the idea that the politicians who reflect these characteristics will "act for" the interests of their group in the issues they raise, policies they champion, and in their wider behaviour in office (Pitkin, 1967). Of central focus is the idea that politicians who share characteristics with a particular group will be best equipped to understand and therefore represent the interests and perspectives of that group (Phillips, 1995). Beyond a very substantial theoretical literature on this topic (see Wolkenstein and Wratil, 2021, for an overview), empirical work has sought to validate the link between descriptive and substantive representation by questioning whether politicians who share descriptive characteristics of particular identities, such as gender, religion, social class, ethnicity, or disability, are more likely to enhance the interests of these

identity groups in legislatures (O'Grady, 2019; Reher, 2022; Sobolewska, McKee and Campbell, 2018).

A natural precursor to addressing this question requires defining what constitutes a “group interest” and determining how politicians might further such interests. Work focusing on this question with respect to gender specifically has highlighted that one way that women politicians substantively represent women is through raising women’s perspectives and experiences when in office. This has been assessed through interviews (Schwindt-Bayer, 2006) and surveys (Wängnerud, 2000) with politicians in a variety of contexts that find women report placing greater emphasis on raising issues related to women, women’s rights, or discrimination targeted towards women. Further, studies have sought to measure whether there are gender differences and have found that, while differences may be conditioned by party or institutional features (Childs and Krook, 2009; Hawkesworth, 2003), women do introduce and sponsor more legislation that serves women’s interest (Swers, 2002), raise women’s perspectives in legislative speeches (Bird, 2005; Clayton, Josefsson and Wang, 2017), and participate more in debates on women’s issues (Bäck, Debus and Müller, 2014).

Beyond raising women’s perspectives, gender differences in the political attitudes and policy priorities of men and women more broadly emerge from socialisation and gender role attitudes (Eagly and Karau, 2002; Eagly, 2013). In reviewing explanations for gender differences, Huddy, Cassese and Lizotte (2008) suggest that differences in political attitudes are related to the personality traits, such as agreeableness, empathy, and assertiveness, that men and women are said to have. Women are said to orientate towards policies that focus on caring for others (Beutel and Marinin, 1995), and “compassion” issues such as poverty, education, children, and health policy issues (Shapiro, 1986). Empirical work has found that, at the party level, the larger numerical presence of women leads to a greater focus on social justice issues (Kittilson, 2011), and to more topically diverse platforms that pay greater attention to ideologically “left” issues (Greene and O'Brien, 2016). This is also reflected in the campaign strategies that women adopt, where they have been shown to pay greater attention to social policy issues than men (Enns-Jedenastik, 2017; Herrnson, Lay and Stokes, 2003). Once in legislatures, this difference has been shown to be present with respect to legislative voting where women have been found to devote more effort to promoting bills that deal with children and family issues, education, and health policy (Schwindt-Bayer, 2006), and are particularly active in advancing policies related to social policy more broadly (Swers, 2002; Thomas, 1994). Further, women tend to frame issues and policies in ways that refer to the individuals specifically affected, such as single mothers or low-income families, whereas men make greater use of abstract orientations such as how they affect the system, state, or economy more broadly (Hargrave and Langengen, 2021).



Differences have also been shown to be present in legislative speeches too, where there is some evidence that women dedicate their speeches to raising issues such as childcare and education, while men focus on foreign policy and infrastructure topics (Boussalis, Mcelroy and Sorace, 2019). Further, while women have overall been shown to participate less than men in legislative debates (Ban et al., 2022), this participation gap has been shown to be exacerbated in debates on “hard” policy issues such as finance or foreign policy (Bäck and Debus, 2019). In the UK specifically, Catalano (2009) examines men and women’s participation in health- and finance-related bills in the 2005-07 parliamentary sessions. She finds that women are significantly more likely than men to participate in debates on health care topics, even when controlling for factors such as party, seniority, or relevant committee membership, however, finds no significant effect of gender on participation in finance related debates.

Gender differences of this sort have also been shown to extend to the portfolios that men and women take on when in office, where women have been shown to be appointed to portfolios that reflect traditional stereotypes (Escobar-Lemmon and Taylor-Robinson, 2005; Krook and O’Brien, 2012), such as in health, social welfare, education, family, and culture, whereas men hold more “masculine” positions such as on economic affairs, defence, employment, and the budget (Barnes and O’Brien, 2018; Blumenau, 2021c; Goodwin, Holden Bates and McKay, 2021). In sum, a wealth of literature across a variety of contexts has shown that women legislators use their time in office to both raise women’s experiences and perspectives and dedicate more effort to social policy issues that are traditionally associated with women.

#### *Explaining the link between descriptive and substantive representation*

Why might men and women politicians differ in the issues they enhance and prioritise in their time in office? Although existing work presents no single unifying theory for *why* these gender differences emerge, I draw on two mechanisms that have commonly been put forward as explanations for this link: shared experiences and (electoral) incentives.

First, perhaps the most prominent explanation is the idea that *shared experiences* among an identity group promotes shared interests and in turn a commitment to promoting these interests in their political roles (Phillips, 1995). The idea is that women have unique experiences that shape and inform the perspectives, priorities, and issues that they will bring to the political arena. Therefore, women politicians, as women, will raise the issues that highlight the challenges and experiences they have encountered. The differences men and women have historically experienced in their everyday lives, such as divergences with respect to caregiving, education, and occupations, divisions of paid and un-

paid labour, and exposure to violence and sexual harassment, will therefore shape the issues they will focus on (Kittilson, 2011; Reingold, 2000). One explanation therefore is that politicians who share descriptive characteristics will be best placed to understand and elevate these perspectives.

Second, the electoral and institutional context can provide strong *incentives* to represent the interests of a particular group. That politicians are motivated by the desire to be reselected, re-elected, and promoted within the party and legislative ranks, and that this motivation will shape their behaviour in office is a long-standing idea (Strøm, 1997). The idea underpinning this mechanism is that politicians are rational actors who will calculate when they stand to benefit most by championing certain issues, or indeed when they stand to incur the most damage by not doing so. The audience for this behaviour can be both voters, who are responsible for electing politicians, and the party leadership, who they rely upon for promotion to higher ranks. Mansbridge (2003) explains that these calculations can often be “anticipatory” in that representatives will behave in a way that they *believe* will benefit them. Gender role theorists content that women are expected to engage in communal behaviours, which are associated with being “warm, gentle, kind and passive”, whereas men are expected to engage in agentic behaviours, which are related to being “tough, aggressive and assertive” (Eagly and Karau, 2002).

Voters have been shown to judge women candidates as better qualified with respect to childcare, welfare, healthcare, education, and “compassion issues”, whereas men are deemed as more competent on issues such as foreign affairs, finance, and defence (Alexander and Andersen, 1993; Huddy and Terkildsen, 1993a). Other work has shown that when the policy domain focuses on masculine issue areas, such as national security or terrorist threat, women candidates are negatively evaluated (Holman, Merolla and Zechmeister, 2011, 2016, 2017). In recent work, Anzia and Bernhard (2022) find women are most successful when they run for elected positions in stereotypically “feminine” offices, such as education boards. Women candidates can also receive more positive evaluations from voters if they emphasise women’s perspectives and issues more broadly (Bauer, 2020b), while incurring punishments if they fail to do this (Cassese and Holman, 2018). The upshot of these arguments is therefore that women politicians face incentives to raise women’s perspectives and traditional “feminine” policy areas to maximise their competence in the eyes of others, and, similarly, to avoid behaviours that may damage them.

#### *Dynamic issue ownership and political experience*

Taken together, the literature suggests that women politicians may be more likely to enhance and raise women’s perspectives, and to dedicate more attention to traditional “feminine” social policy ar-

eas. However, few studies have considered how these dynamics change over the course of politicians' careers. I build on the perspectives described above that gender differences arise because of both women's shared experiences in, and therefore better understanding of, the issues that affect women, but also because of the incentive to appear competent and credible in the sets of issues they raise. Building on this logic, I argue that gender differences in the extent to which women politicians raise issues stereotypically associated with women will be most pronounced when they are junior and inexperienced, however, this difference will decline as women spend more time in office and gain more experience and are therefore able to diversify the sets of issues they raise beyond those determined by their gender.

First, when new to legislatures, politicians undergo a process where they learn how to best succeed in their careers (Rush and Giddings, 2011). In the early stages, politicians are at their most inexperienced and will therefore gravitate towards political issues where they have prior expertise and interest, as well as strategic considerations relating to the desire to appear competent and credible (Bailer and Ohmura, 2018). As Bailer et al. (2022, 4) argue, parliamentarians will "engage in activities that they can credibly represent and in which they have personal experience and thus expertise". While the desire to appear competent is present for both men and women, these early career calculations may be particularly acute for women given that legislatures are traditionally male-dominated institutions that favour and re-create masculine norms and practices (Krook and Mackay, 2011). Further, women face a "burden of doubt" from those who traditionally belong in the space to prove their competency and justify their presence (Puwar, 2004) and have to "navigate an electoral arena which is much more precarious for them than it is for men" (Lazarus and Steigerwalt, 2018, 172). Women will likely raise issues that are traditionally associated with women because, first, they have more genuine expertise in these areas by virtue of their shared experiences with women, and second, they may be attributed more experience, credibility, and authority in these areas because of the stereotypical associations between social policy issues and women. While men will also gravitate towards issues where they have prior expertise, these are unlikely to be determined by their gender specifically to the same extent as women.

However, as they spend more time in office, they learn to navigate their roles, and become more experienced, effective, and efficient in their work (Rush and Giddings, 2011). As Bailer and Ohmura (2018, 496) argue, "expertise is nearly automatically gathered over time in parliament and is the result of a necessary learning process". Over time, politicians will participate in a variety of debates, will ask questions, sit on committees, be lobbied on policies and campaigns, and engage with their constituents' concerns. In doing so, they gain greater insights and experience in a wider variety of issues

and will therefore both naturally diversify the set of policy areas they dedicate their attention towards, but also strengthen their portfolios of expertise. Indeed, [Miquel and Snyder \(2006, 347\)](#) find that the effectiveness of legislators increases sharply with increased tenure in parliament, primarily through “learning by doing”. Therefore, as a function of their increased experience in office, the sets of issues that women raise will be determined by factors beyond only their gender, and instead by a wider variety of factors too. While junior women parliamentarians therefore dedicate significant efforts to raise issues they are stereotypically expected to know more about, the selection of issues will diversify as they spend more time in parliament and gain more expertise and credibility in their roles.

Second, women politicians may also internalise the expectations from voters that they are responsible for and may be rewarded for raising women’s perspectives and traditional social policy areas more broadly ([Anzia and Berry, 2011](#)). Women face external pressures from voters that the furthering of women’s interests in the political realm is an expected part of their role, and are “blamed if they do not pursue ‘women’s issues’” ([Bergqvist, Bjarnegård and Zetterberg, 2018, 581](#)). Associatedly, descriptive representatives overall ([Collet, 2008](#)), but women candidates and politicians in particular, are ascribed more authority when raising issues relevant to their identities. Candidates are incentivised to pander to voters on the sets of policies where they are likely to have a prior advantage ([Adams, Merrill III and Grofman, 2005](#)), which, for women, has been shown to be “feminine” issues which they are perceived as better suited to ([Herrnson, Lay and Stokes, 2003](#); [Huddy and Terkildsen, 1993a](#)). The implication of these arguments are that women politicians dedicate more efforts to raising women’s perspectives and traditional “feminine” policy areas because voters perceive them as primarily responsible for these efforts and will ascribe them more competence and authority on these issues.

However, just as politicians gain status in the eyes of their colleagues as their political careers progress, they also gain credibility from voters. A wealth of work has shown that voters place a higher premium on the experience in office that incumbent politicians gather ([Eggers and Spirling, 2017](#); [Fulton, 2012](#)), that this perceived experience results in politicians being far more likely to be re-elected in a wide variety of contexts ([Saha and Weeks, 2022](#); [Schwarz and Coppock, 2022](#)). While work has shown women politicians face more significant hurdles than men when they first run for office, this gap has been found to be smaller once they have entered office and decide to run again ([Huddy and Terkildsen, 1993b](#); [Kahn, 1992](#)). Further, experimental work assessing how voters judge politicians’ past performance has shown voters to be no more critical of and no less likely to re-elect women than men ([de Geus et al., 2021](#); [Hargrave and Smith, 2022](#)). Therefore, with the overall increased perceived credibility that comes with experience, women need to rely less on focusing on issues where they are

traditionally expected to have authority and focus on a wider diversity of issue areas. These arguments suggest that women may raise traditional women's issues at the beginning of their careers to capitalise on the areas where they are likely to be perceived as having the most expertise, however, the need to do this will likely be eclipsed by the experience they gain the longer they serve in office.

In sum, in the early stages of their careers, when most inexperienced, politicians will gravitate towards issues where they both have the most prior experience and, crucially, are perceived to have expertise by others. For women, this is likely to be issues traditionally associated with women. However, as they spend more time in office, learn to navigate their roles, establish themselves in other policy areas, and gain more experience and credibility, the attention given to this narrower set of issues will likely decline. Of course, experienced women are unlikely to cease to focus on these issues entirely, but rather they will also focus on a wider set of issues and go on to raise women's perspectives and experiences with respect to these issues more broadly. Taken together, my central argument is that women politicians will focus more on traditional women's issues than men, but that this gap will be largest among junior politicians and will diminish with increased experience.

Before proceeding, I note one key study that shares similarity with my own. Recent work by [Bailer et al. \(2022\)](#) examines the (changing) relationship between descriptive and substantive representation over politicians' careers in the German Bundestag. Focusing on legislators from four historically disadvantaged groups – women, migrants, low social class, and the young – they argue that, as their careers progress, they will strategically decide to reduce the substantive representational efforts of their group in favour of progressing up the hierarchy of legislative roles. They find evidence consistent with this argument for migrants, low social class, and young legislators, but no such effect for women. While this paper shares notable similarities with my own, my work makes progress in several ways. First, as I focus solely on how these dynamics operate for women, my theoretical argument differs notably from that of [Bailer et al. \(2022\)](#), as I describe how the pressure for women to focus on stereotypically “feminine” policy areas decline over the course of their careers. In doing so, my argument draws on a wider range of literature on gender stereotypes, electoral bias, and incumbency advantage to more fully outline how these pressures operate for women specifically.

Second, [Bailer et al. \(2022\)](#) focus only on the extent to which legislators raise arguments that relate directly to their group. For instance, the extent to which women legislators refer to women or migrant politicians to migrants. Instead, I focus on a wide range of policy areas which are central in contemporary British politics. Third, I focus both on the extent to which women politicians dedicate more attention to raising stereotypically “feminine” policy areas *and* the extent to which women raise

women's experiences in debates across the policy spectrum.

### 4.3 Data and methodology

In the UK, policies are most commonly proposed centrally by the government and MPs commonly vote in accordance with their party's position (Cowley and Childs, 2003). Therefore, MPs' voting behaviour is not an appropriate place to examine how politicians prioritise issues as this is unlikely to reflect individual preferences. Instead, I use politicians' speeches to identify the issues they individually engage with. Parliamentary speeches represent a useful source for identifying individual priorities over the course of politicians' careers and are now a commonly used source of information for measuring politicians' behaviour with respect to both style (Dietrich, Hayes and O'Brien, 2019) and issues (Blumenau and Damiani, 2021). I study UK House of Commons debates between May 1997 and March 2019, using data collected and made available by Blumenau (2021a). For all measures of issues I outline below, I collapse the data such that all speeches made by an MP in a debate constitute a single speech-document, making my unit of analysis an individual in a debate. The sample consists of 23,620 debates, 437,503 MP-debate observations, and 1,422 MPs in total (370 women, and 1,052 men).

#### *Conceptualising and measuring topic areas*

My key substantive interest is to empirically examine whether there are gender differences in the sets of issues that politicians raise. This includes both issues that the work described above has highlighted as stereotypically associated with women or men, and a wider range of policy areas that have been identified as central to British politics according to the UK Policy Agendas Project (Bevan and Jennings, 2019). These are agriculture, children and family, crime and policing, defence, education, environment, finance and economy, health, trade, transport, and welfare. To identify the extent to which politicians talk about these issues, I apply the same word-embedding dictionary-based approach described in chapter 3. This approach combines dictionaries with a locally-trained word-embedding model to construct domain-specific dictionaries that are intended to capture the issues as they manifest in the specific context of UK parliamentary debate. A common criticism of off-the-shelf dictionary approaches is that words used in one context may be different to words used in another context (Grimmer and Stewart, 2013). Instead, the word-embedding approach enables me to create a series of bespoke dictionaries that are firmly rooted in the way that issues are discussed in parliamentary debate. The approach has been shown to far outperform traditional dictionary approaches and is of increasing use

in political science applications (Osnabrügge, Hobolt and Rodon, 2021; Rice and Zorn, 2021; Rodriguez and Spirling, 2022; Widmann and Wich, 2022).

For each issue area, I replicate the measurement steps outlined in chapter 3. A key component of this approach requires the defining of a “seed” dictionary that represents the concepts of interest. I select these words through a combination of extracting the key sub-areas listed in the Policy Agendas Project (Bevan and Jennings, 2019) and reading parliamentary debates on relevant areas and selecting words and phrases that were likely to capture these issue areas. Table B.1 in section B.1 in the appendix shows the seed dictionaries for each area. This provides me with a measure of the extent to which each individual MP in a given debate is speaking about the issue areas. I normalise each measure across documents to have mean zero and standard deviation one. This means the average gender gaps across the various issue areas can be interpreted in standard deviations of the outcome variables.

### *Validation*

In section B.2 in the appendix, I carry out a series of validation exercises to assess the performance of the word-embedding based approach. Alongside face validation of the top words (table B.2) and sentences (tables B.3–B.6) associated with the policy areas, I also assess, first, whether politicians who hold relevant ministerial positions or sit on relevant committees talk more about a given issue area. For instance, whether a politician who holds a health ministerial position talks more about health than a politician who does not. The results from the validation (figure B.1) are reassuring, and show that, for all issue areas, politicians with relevant ministerial or committee assignments do talk significantly more about their respective issue areas.

Second, the word-embedding based approach provides me with a measure of the degree to which each individual contribution by an MP in a given debate discusses each of the issue areas. If the approach is working as intended, then we would expect that speeches that pay particular attention to certain issues are nested in debates on the same issues. For example, whether speeches that include many words associated with welfare are delivered in debates that are also on welfare issues more broadly. I assess this both qualitatively and quantitatively. Qualitatively, I aggregate the word-embedding scores from the speech-level to the debate-level to identify the degree to which a debate overall is about a given issue area. I then extract the titles of the top 20 scoring debates for each issue area. For all issue areas, the results (tables B.7–B.10) are reassuring. For instance, the finance and economy topic includes debates on taxation, income, and benefits. Similarly, debates on agriculture refer to issues such as food labelling, animal welfare, and farm produce. Therefore, qualitatively at

least, debates that contain many speeches on a given issue area seem to be nested in well-matched debates topically.

I also assess the debate-level validity of the issue scores quantitatively. To do so, I apply the words from table B.1 as dictionary measures to the debate titles. I identify the number of words from each topic dictionary that appears in a given debate title and classify the debate as about the topic area for which the most words appear. This provides me with a measure of the topic of each debate. I then assess whether MPs deliver more speeches on the topic area that relates to the debate topic. Again, reassuringly, in figure B.2, I show that this is true for all issue areas. More detail on each of the validation exercises and the full results are presented in section B.2 in the appendix.

#### 4.4 Empirical strategy

My goal is to assess the degree to which there are gender differences in the topics MPs raise in parliament, and how these change with increased political experience. To investigate these quantities of interest, I estimate a series of OLS regression models that allow me to account for a host of individual-level characteristics, described below, that may differ between men and women and be correlated with the sorts of issues that MPs raise in debates. To assess these questions, I estimate a series of models for the topic-based outcomes  $Y_{id}^t$  for an individual MP  $i$  in a debate  $d$  for topic  $t$  of the following form:

$$Y_{id}^t = \alpha + \beta_1 \text{WomanMP}_i + \sum_{k=1}^K \gamma_k X_{ik} + \epsilon_{id} \quad (4.1)$$

where  $\beta_1$ , the main quantity of interest, describes the average difference between men and women with respect to  $Y_{id}^t$ , while controlling for various time-varying MP-level covariates. I include several such controls. First, a likely explanation for the sorts of issues MPs raise in debates is the party they represent. In the UK, the Labour Party has historically had proportionally more women than other parties, and we therefore might be concerned that men and women will raise different issues simply as a function of the parties they represent. To address this concern, I include a set of party dummies.

Second, with increased time spent in parliament, MPs are more likely to take on new committee chair or ministerial responsibilities. By holding a given committee position, it seems likely that MPs will have greater expertise on, or interest in, a given issue and will therefore bring this into the debates they participate in. Further, MPs that hold frontbench positions for either the government or opposition act as spokespeople for particular issue areas. MPs with these positions will therefore not



only disproportionately participate in particular issue areas, but parliamentary convention dictates that they will also tend not to speak on debates outside of these areas. To account for MPs taking on new issue-specific roles the longer they are in parliament, I also control for whether an MP holds a frontbench position for either the government or opposition at the time of speech delivery, as well as whether they are a committee chair. Third, over the period I study, the party of government switched from the Labour Party in 2010, who has proportionally more women than other parties, to the Conservatives. To account for the potential of government-opposition dynamics, I control for whether an MP is a member of a governing or an opposition party in each period. Finally, recent work has emphasised that electoral vulnerability (Hömann and Nugent, 2021) determines the extent to which men and women MPs raise certain issues in debates. In particular, men have been shown to raise women’s issues to a greater extent when they represent marginal constituencies. I control for the percentage-point margin of victory of MPs in the previous election.

Next, as argued above, the main substantive quantity of interest here is the extent to which gender differences in the issues politicians discuss change with increased parliamentary experience. To assess whether the relationship between time in parliament and the issues raised is different for men and women MPs, I estimate:

$$Y_{id}^t = \alpha + \beta_1 WomanMP_i + \sum_{l=2}^L \varphi_l^1 YearsinParliament_{i,l} + \sum_{l=2}^L \varphi_l^2 (WomanMP_i \cdot YearsinParliament_{i,l}) + \sum_{k=1}^K \gamma_k X_{i,k} + \epsilon_{idt} \quad (4.2)$$

where  $YearsinParliament$  is a matrix of dummy variables with cells which take the value of 1 when MP  $i$  has spent  $l$  years in parliament at the time of delivering a given speech.  $\varphi_l^1$  therefore captures the effect of spending  $l$  years in parliament (relative to an MP being in their first year in parliament) on the attention given to topic  $t$  for men MPs, and  $\varphi_l^2$  describes the difference in this effect for women MPs. Accordingly,  $\varphi_l^1 + \varphi_l^2$  describes the effect of spending  $l$  years in parliament (compared to one year) on the outcome for women MPs. Again,  $X_{ik}$  is the same set of individual-level controls described above. Below, I plot the difference between men and women,  $\varphi_l^2$ , in figure 4.2 and the average topical usage for men and women,  $\varphi_l^1$  and  $\varphi_l^1 + \varphi_l^2$ , in figure 4.3.

Finally, my argument suggests that MPs change the topics they speak about the longer they are in parliament because of their increased levels of expertise, experience, and credibility in a wider variety of issue areas the longer they serve. However, an alternative explanation may be that certain kinds

of politicians serve longer parliamentary careers, and that the women who stay for longer happen to talk less about “feminine” policy areas, whereas the women who serve only a single term talk more about “feminine” areas and then leave parliament. Further, it seems likely that there will be common trends in the prevalence of the topical outcomes over the period studied. To account for both potential source of confounding, I estimate a two-way fixed effect model of the following form:

$$Y_{idt}^t = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \sum_{k=1}^K \gamma_k X_{ik} + \lambda_i + \delta_t + \epsilon_i \quad (4.3)$$

where  $X_{ik}$  represents a set of time-varying MP-level covariates (cabinet membership, shadow cabinet measurement, government or opposition, minister, shadow minister, committee chairs, and marginality),  $\lambda_i$  represents MP fixed effects which allow me to account for any omitted variables that vary across MPs but are constant within MPs over time, and  $\delta_t$  represents time fixed effects for parliamentary terms, of which there are 6 between 1997 and 2019, allowing me to account for any omitted variables that change over time for all politicians. I cluster standard errors in all models at the MP level  $i$ .

## 4.5 Results

Do men and women raise different issues in parliamentary debate? In figure 4.1, I present the results from the model described in equation 4.1. Positive values indicate that women talk more about a given topic, and negative values indicate men talk more about a given topic. The point estimates and confidence are coloured in black if the gender difference is associated with a  $p$ -value of less than 0.05, and otherwise in grey. The figure shows that, at least at the aggregate level, the topics that men and women raise are broadly consistent with traditional divisions highlighted in prior literature. Women talk more about children and family, health, women, education, and welfare – all stereotypically “feminine” policy areas – as well as talking more about crime and policing. Similarly, men talk more about defence. For the environment, transport, finance and economy, trade, and agriculture I uncover no statistically significant gender differences.

How does political experience affect the issues men and women raise in parliament? In figure 4.2, I show how gender differences in topics change with increased time in parliament. For some issue areas – agriculture, transport, finance and economy, and trade – there are near consistent statistically insignificant differences between men and women at all points of parliamentary seniority. Women

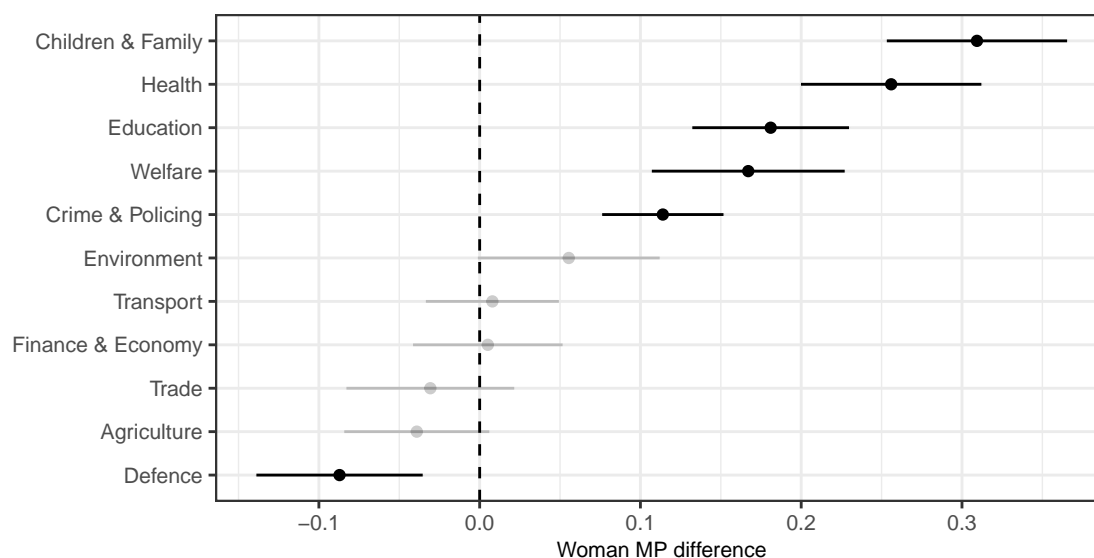


Figure 4.1: Gender differences in issue areas

talk significantly more about crime and policing than men at almost all points, except from around 16 years onwards where the difference is no longer significant. For other areas, such as defence and the environment, the size of the gender gap changes, but follows no clear trend.

Figure 4.2 also shows that there is significant variation in the size of the gender differences for all social policy areas. For children and family, health, education, and welfare, there are large gender differences between junior men and women, however these gaps decrease markedly as time in parliament increases. Recall, the topic-based measures are standardised to have mean zero, standard deviation one, and the size of these changes are therefore non-trivial. Focusing on the point estimates for education, with only one year of parliamentary experience, women MPs speak 0.27 standard deviations more about education than men, however this decreases to only 0.13 standard deviations among politicians with eight years of parliamentary experience.

Further, by approximately 15 years in parliament, there are no longer statistically significant differences between men and women. However, the width of the confidence intervals for the point estimates representing gender gaps at this point of parliamentary tenure are wide. This is because very few women in the data have this level of parliamentary experience. In fact, of the 370 individual women MPs in the data, only 36 have experience in parliament of 20 or more years, compared to 313 men out of a total of 1,052.<sup>1</sup> Realistically, there is simply insufficient data to make precise statements at this

<sup>1</sup>Of the MPs in the data, 844 men and 218 women have at least 5 years of experience, 634 men and 128 women have at least 10 years of experience, 454 men and 54 women have at least 15 years of experience, and 313 men and 36 women have at least 20 years of experience in parliament.

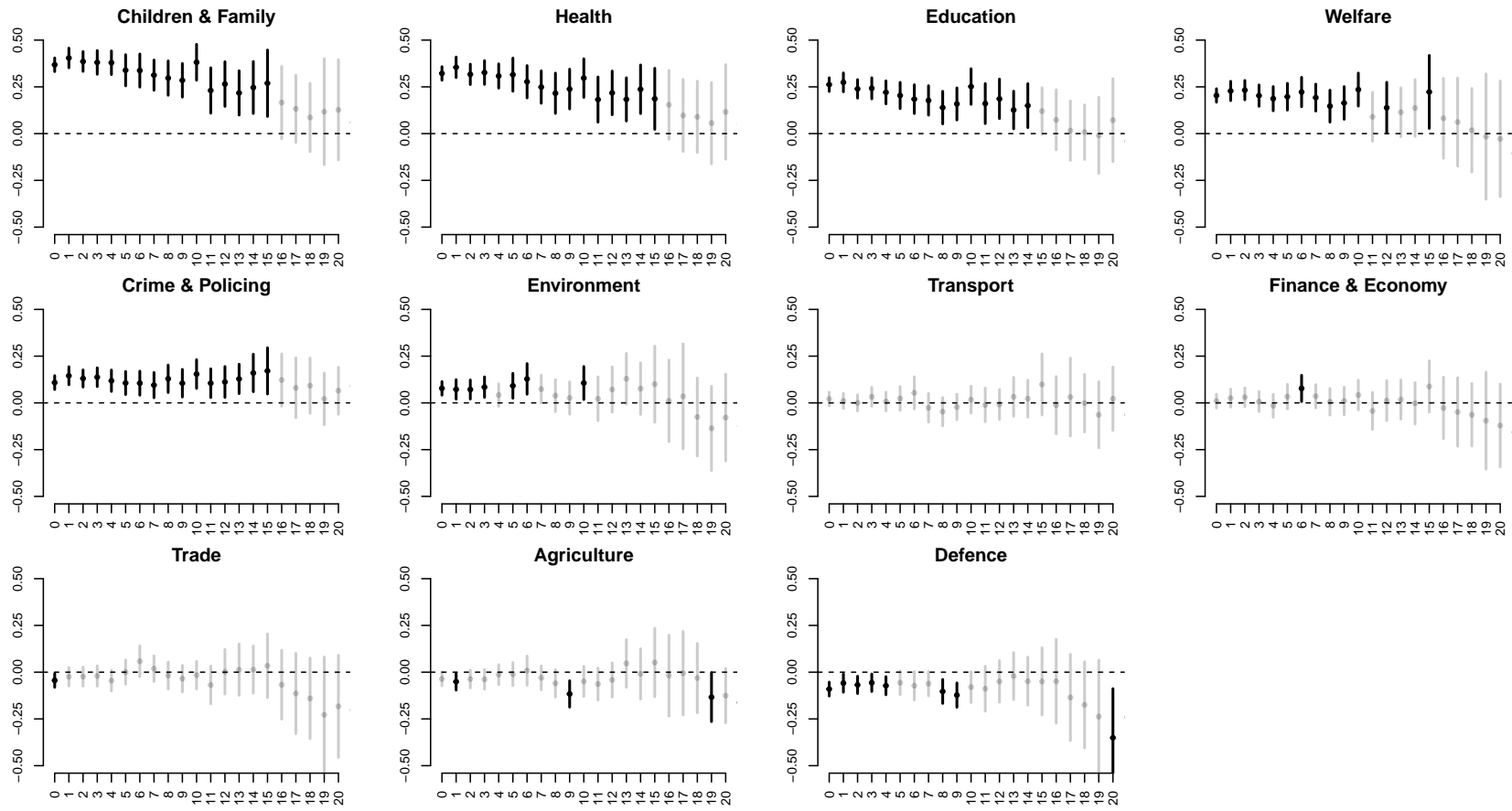


Figure 4.2: Gender gaps in issues over parliamentary tenure

point in a parliamentary career.

Nonetheless, the trends for the traditional “feminine” social policy areas clearly follow the main argument outlined above: that junior women parliamentarians will focus more on issues traditionally associated with women, but that this will decrease with increased parliamentary experience. Notably, senior women do not entirely cease to focus more on these social policy areas than men, which lends support to the perspectives introduced above that shared experiences will also inform the sets of issues that women raise. However, I find clear evidence that the difference is most pronounced in the early stages of women’s careers and declines as they gain more experience.

Are the changing gender gaps in the discussion of topic areas the result of changes in men or women’s behaviour? In figure 4.3, I show the changes in topic prevalence for men and women separately as years in parliament increases. For many of the issue areas – environment, transport, finance and economy, trade, agriculture, and defence – there are near consistent statistically insignificant differences between men and women at all points of tenure. For crime and policing, men and women remain relatively stable throughout their time in parliament, and women near consistently talk more about this topic than men. The figure also lends further support for my central argument and the findings presented above: while there is an overall decrease in the extent to which all politicians talk about the traditional “feminine” social policy areas, this decrease is notably steeper for women than it is for men.

A puzzle that emerges in figure 4.3 is that, across almost all issue areas, there is a downward trajectory for men and women alike. On which issue areas do senior politicians dedicate their speeches to raising? The word-embedding based approach that I use to measure the topic areas of interest does not offer a complete decomposition of the topical agenda in parliament, but rather is a supervised method guided by the set of topics that are specified. Other common approaches to measuring the topical focus of political speech, such as structural topic models, are unsupervised and do however capture the entire population of topics that MPs raise in their speeches. In the appendix in section B.4, I use statistical topic models to identify a wide variety of topics that politicians raise and identify the sets of topics that *increase* with parliamentary seniority. Intuitively, I find that the topics which increase in prevalence with time in parliament are topics which are not captured by my word-embedding based approach. In particular, the largest positive effects relate to topics that are procedural in nature: such as references to parliamentary address terms, parliamentary process, or the legislative process. The upshot from this analysis is that as politicians spend more time in parliament, they seem to dedicate increasing attention to matters of procedure and process relative to junior politicians.

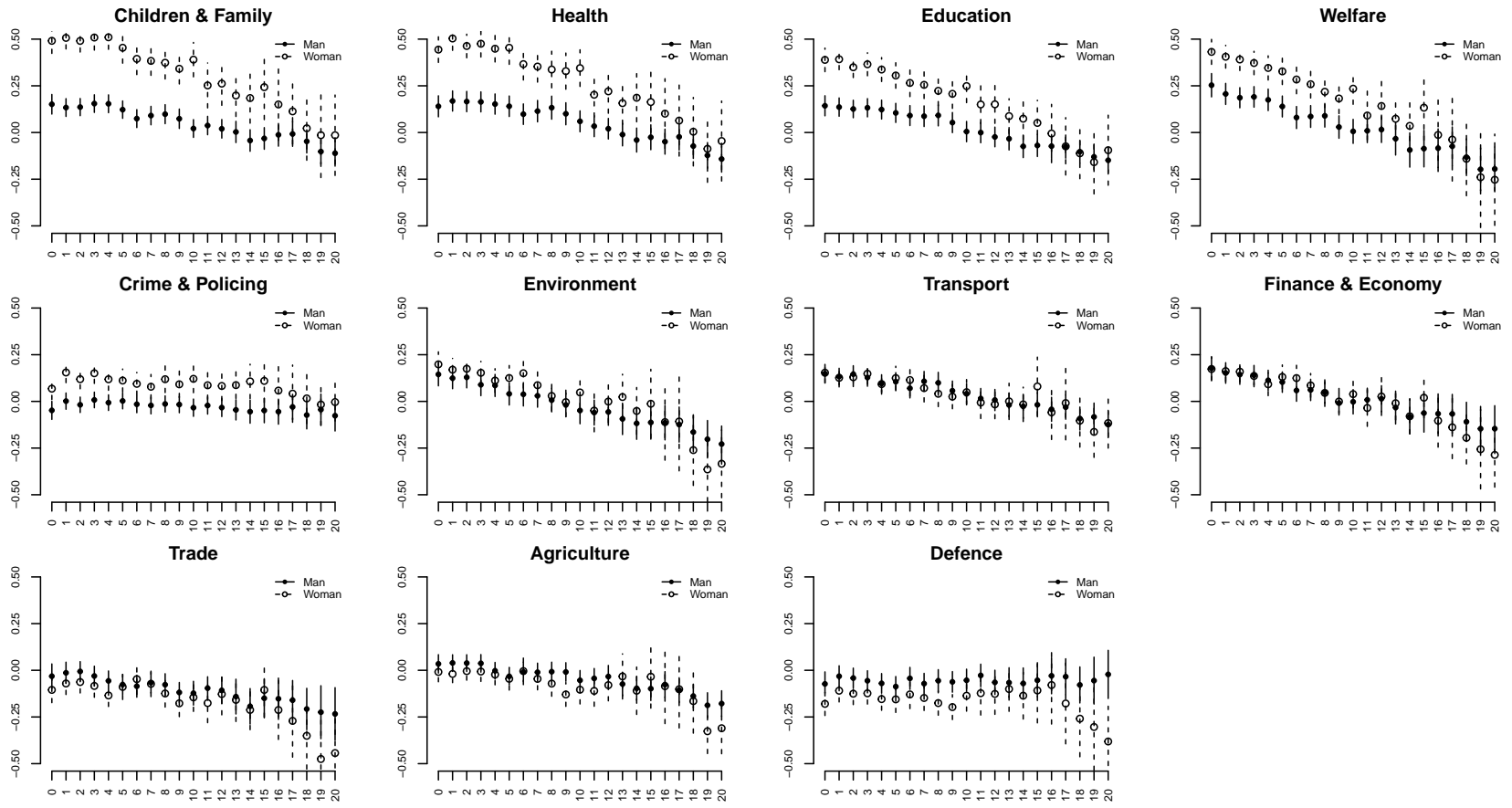


Figure 4.3: Average topic use for men and women over parliamentary tenure

## 4.6 Amplifying women's experiences across the policy process

Do women politicians raise women's concerns, experiences, and perspectives across the policy process? In the previous sections, I have addressed one way in which men and women are said to differ: that they raise substantively different political issues. However, as described previously, work on descriptive and substantive representation also poses that gender differences are said to arise because of women's *shared experiences* in and therefore better understanding of the issues that affect women (Phillips, 1995). In this section, I examine whether women politicians do more to represent women in their parliamentary speeches, how this changes across topic areas, and throughout the course of their careers.

To do so, I identify whether speeches include many direct references to women. I focus on speeches that include many words that fall into the Linguistic Inquiry and Word Count (LIWC) dictionary's *Female* category (Pennebaker et al., 2015). This includes words such as "woman", "women's", "female", "female's", "girl", "girl's", among others. The application of this approach is simple. I apply the LIWC Female category to the speech-documents, identify the number of words from the dictionary that appear, and then calculate these words as a proportion of all words spoken in a given MP-debate contribution. To put this measurement approach on a comparable scale to the word-embedding based approach, I normalise each measure across documents to have mean zero and standard deviation one. Although somewhat crude, this approach has been used in previous studies (Pearson and Dancey, 2011b), and has been shown to correlate highly with more sophisticated approaches (Dietrich, Hayes and O'Brien, 2019). Further, I opt to use the simple dictionary-based approach here as opposed to the more sophisticated word-embedding approach described above because my interest is explicitly in the extent to which MPs make direct references to women in their speeches, and not in developing a topic related to "women" more broadly which the word-embedding model would detect.

With this measure in hand, I assess whether there are gender differences in the extent to which politicians raise women's perspectives over the course of their careers. In figure 4.4, I replicate the analysis presented in figures 4.2 and 4.3 above. The left panel shows the gender gaps among men and women with increasing levels of seniority and shows that women make significantly more references to women than men, and that this gender difference remains relatively stable. This finding departs from the trends described above, but tells a potentially positive story, as it suggests that while women are able to diversify the sets of topics they raise the longer they are in parliament, they continue to raise women's perspectives. The right panel shows the average references to women by men and women

politicians separately. I find that women MPs consistently refer to women more than men, and that this difference, again, remains relatively stable throughout their parliamentary careers. Both figures, therefore, lend support to the idea that women politicians raise women's perspectives more than men, and that they continue to do so throughout the career cycle.

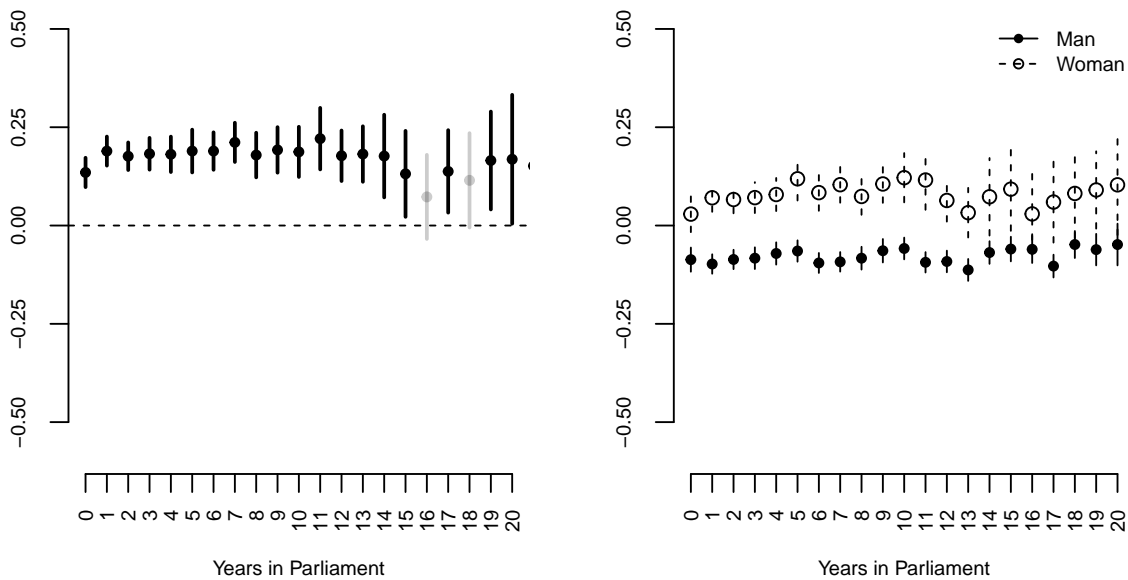


Figure 4.4: Effect of time in parliament on gender gaps in raising women's perspectives

Do women raise women's perspectives more than men in debates on all topic areas? One explanation for the findings I present above may be that men and women systematically participate in different kinds of debates. For instance, women may disproportionately participate in debates on children and family, and these debates include more frequent references to women's experiences. Alternatively, it may be the case that there are not differential participation rates by topic, but that women politicians raise women's perspectives across a wide variety of topical areas. In figure 4.5, I disentangle these explanations by reporting the results from an analysis where I identify the topic of each debate, subset to debates on each topic individually, and identify whether there are gender differences in the extent to which politicians refer to women across these issue areas.

The y-axis in figure 4.5 denotes the topic of the debate and the x-axis shows the average gender difference in debates characterised by each topic area. Positive values indicate that women MPs talk more about women and negative values would indicate that men talk more about women. As is clear from the figure, there are no negative values, and women MPs therefore raise women's perspectives



consistently across debates on a wide variety of topics. The magnitude of the coefficients does, however, vary across debate topics. The smallest differences are on health and agriculture, while some of the largest differences between are in “masculine” policy areas such as crime and policing, trade, and defence. This suggests that women bring women’s perspectives when participating on these issue areas.

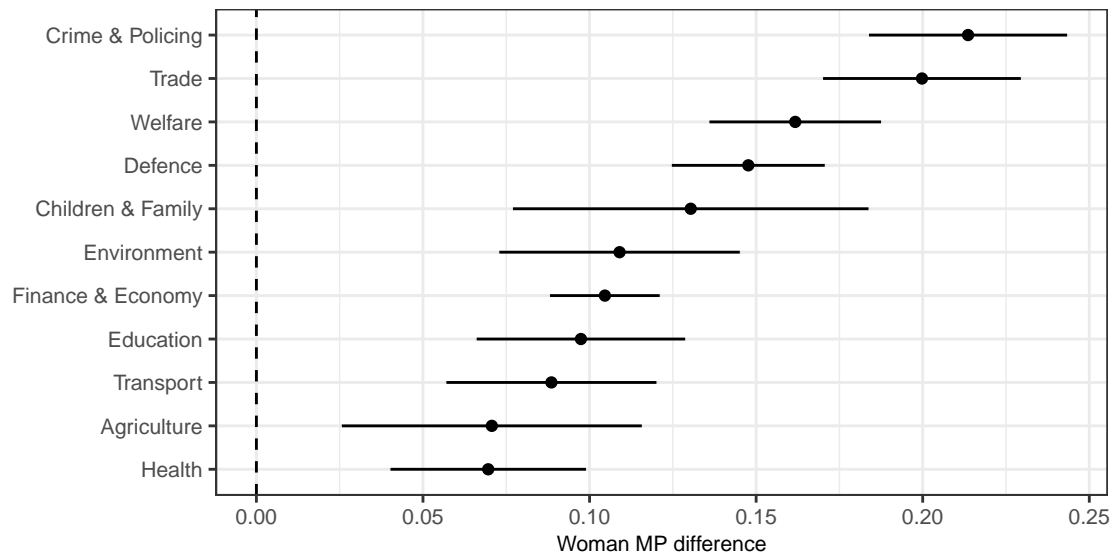


Figure 4.5: Gender differences in raising women’s perspectives across debates on different topics

Are these differences consistent with increased seniority? In the appendix, in tables [B.23–B.24](#), I present the results from an analysis assessing whether this is the case. I estimate a series of models where the outcome is the measure of references to women, and the explanatory variables are MP gender, years in parliament, and the interaction between the two. This enables me to identify whether the effect of increased time in parliament on the propensity for politicians to raise women’s perspectives is different for men and women. I find that, for all topic areas, women politicians raise women’s perspectives more than men and that this difference is unaffected by political experience. The implication of these findings is positive, as they suggest that across debates on a wider variety of topics, women politicians raise women’s perspectives and they continue to do this even with increased parliamentary experience. Further, some of the largest effects are concentrated among the more traditionally “masculine” policy areas, such as crime and policing or defence. This suggests that while women amplify the experiences of women in debates across the policy spectrum, they dedicate particular attention to doing so in the more traditionally “masculine” policy domains.

*Alternative explanations*

I have argued that in the early stages of their careers, women politicians will likely dedicate significant efforts to raising issues that are traditionally associated with women due to the desire to seem competent and credible. However, as they spend more time in office, this pressure will likely decline. There are, however, alternative explanations that might explain the patterns I document.

First, it is possible that part of the motivation for why politicians enter institutions is likely to be because of a desire to bring about change for the sets of issues they care passionately about. Indeed, in interviews with women, many reported their motivation to bring about positive change (Sones, Moran and Lovenduski, 2005, 136). Therefore, a rather pessimistic alternative explanation for why there is a decline in the degree to which women talk about social policy issues throughout their careers might be that they enter parliament, try and push for positive change on these issues, are unsuccessful in bringing about change on this front, and decide to instead move onto issues where they are likely to have more success. For this explanation to explain the trends I document, it seems likely women would also decrease the degree to which they raise women's perspectives. I see no evidence that this is the case, but rather that they do this more than men, do so in debates on each of the topical areas, and that this does not decline with increased seniority. This alternative motivation for behaviour change seems unlikely to explain the patterns I document.

Second, we might be concerned that a potential explanation for the changes is that, with increased time in parliament, MPs are likely to take on ministerial and committee responsibilities and that these leadership positions will determine the issues they focus on. In the UK, MPs that hold frontbench positions for the government or opposition parties act as spokespeople for particular issues. MPs with these positions will therefore participate only in debates specific to their issue areas and will not participate in debates on issues more broadly. This may be of particular concern given the evidence that men and women MPs in the UK and further afield hold systematically different ministerial positions (Blumenau, 2021c) and committee assignments (Goodwin, Holden Bates and Mckay, 2021). Therefore, the changes shown here could be a simple reflection of parliamentary career progression into ministerial positions. However, recall that the models presented in figures 4.2–4.3 control for a host of institutional covariates such as whether an MP attends (shadow) cabinet, holds a committee chair position, or holds a ministerial position for either the government or opposition at the time of speech delivery. Therefore, even accounting for the fact that politicians will likely take on ministerial positions in potentially a variety of policy areas the longer they stay in parliament, I still find that the amount

of attention women dedicate to social policy issues, but not issues more broadly, declines. Therefore, the trends documented above cannot be explained by MPs taking on ministerial or committee responsibilities as their careers progress.<sup>23</sup>

Third, certain kinds of politicians serve longer parliamentary careers, and the women who stay in parliament for longer happen to talk less about traditional “feminine” policy areas, whereas the women MPs who serve only a single term talk more about “feminine” policy areas and then leave parliament. While my story suggests that within-MP change is responsible, the results I present above do not enable me to disentangle which of these explanations – within-MP or selection effects – may be responsible for the aggregate change. Further, it seems likely that over the period under study, there may be events that affect the prominence of particular issues for all MPs. Imagine, for instance, that salient events have occurred that relate to violence against women throughout the period studied. This would likely increase the prominence of the crime and policing topic in politicians’ speeches and be a source of confounding if women in particular increased their attention to this topical area. To account for these potential sources of confounding, I estimate the model described in equation 4.3, and assess the effect of years in parliament on the topic-based outcomes and include MP and parliamentary term fixed effects.

In the appendix, I present the full regression outputs in tables B.11–B.22, however, for ease of interpretation, I present the results separately for men (in black solid points) and women MPs (in grey hollow points) in figure 4.6. The figure shows the average topical prevalence for men and women while accounting for time-varying MP-level covariates, MP fixed effects, and parliamentary term fixed effects. There are several things to note. First, for most topic areas, although both men and women talk less about them with increased time in parliament, there is no significant difference *between* men and women. Second, the largest gender differences I observe are for many of the “feminine” social policy areas. Notably, the only areas for which I observe a significance difference between men and

<sup>2</sup>In tables B.11–B.22 the appendix, I present the results from an analysis estimating a series of models with “strict” controls for ministerial or committee positions. For each issue area, I create a set of binary variables for whether an MP holds a government or opposition ministerial position, or sits on a committee, that is relevant to each of my issue areas. For instance, whether an MP holds a ministerial position for health, education, or trade, or sits on a transport, environment, or defence committee. I then add these issue-specific controls to the relevant models. I find that the results I present here hold even with these stricter controls.

<sup>3</sup>In the appendix in section B.5, I estimate five sets of models for each topic area with increasingly conservative specifications. For each, the outcomes are the issue areas. First, I interact MP gender and years in parliament. Second, I interact MP gender and years in parliament and also add a series of general individual-level controls (party, government or opposition status, cabinet membership, shadow cabinet measurement, minister, shadow minister, committee chairs, and marginality). Third, I interact MP gender and years in parliament and also add a series of *specific* controls for the relevant ministerial or committee positions and a series of general individual-level controls (party and marginality). Fourth, I interact MP gender and years in parliament, control for a series of general individual-level controls, and include MP fixed effects. Finally, I interact MP gender and years in parliament, control for a series of general individual-level controls, include MP fixed effects, and fixed effects for each parliamentary term. I find that the main results I present here are robust to these more conservative estimations.

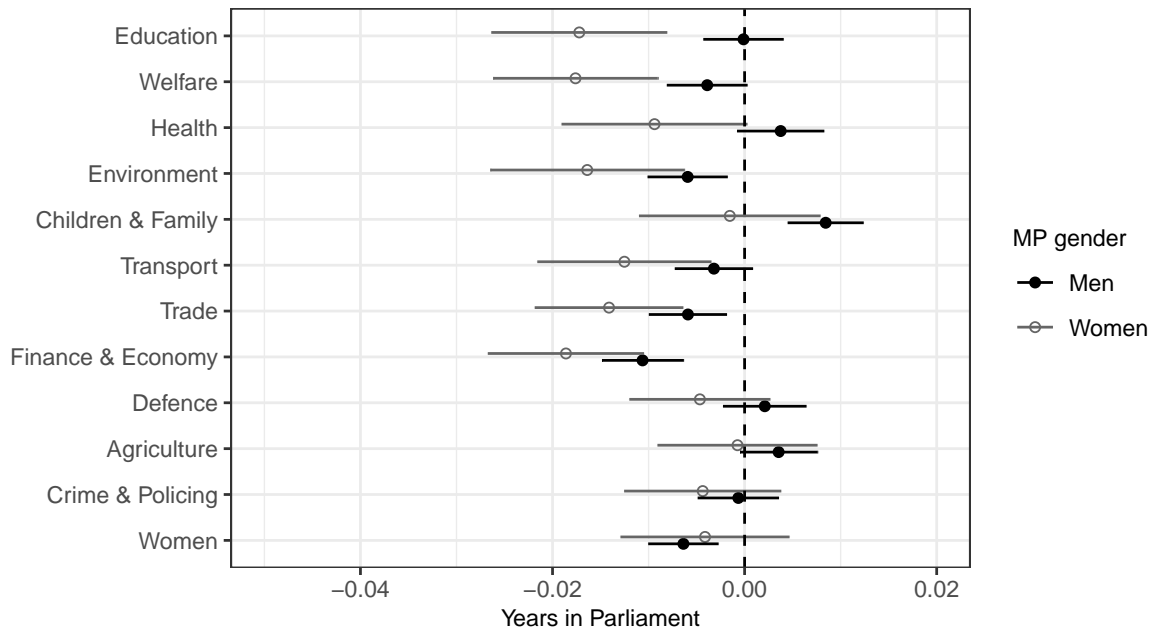


Figure 4.6: **Effect of years in parliament on topic discussion for men and women:** The figure shows a two-way fixed effect specification with MP and parliamentary term fixed effects.

women are for these areas. While men therefore talk less about these policy areas as their careers progress, women talk about them even less. The patterns shown here are consistent with the story from figure 4.2: even when accounting for any omitted variables that vary across MPs but are constant within MPs over time *and* any factors that change over time for all MPs, the attention that women MPs dedicate to “feminine” policy areas, relative to men, decreases as they become more senior in parliament. Fourth, while I find that the inclusion of MP and term fixed effects shows that men talk less about women the longer they are in parliament, there is still no evidence that women decrease their focus on representing women. This finding is consistent with the variety of results I have presented above that women politicians continue to raise women’s perspectives throughout their careers.

## 4.7 Conclusion

How does political experience affect the issues men and women raise in parliamentary debate? Past work has emphasised that men and women politicians speak about systematically different sets of political issues, as women are expected to focus more on “feminine” social policy issues that reflect their traditional gender roles as caregivers (Norris, 1996), and the wider stereotype that women are

more communal, caring, and compassionate (Eagly and Karau, 2002). Focusing on the UK, I argue that while junior women parliamentarians who lack experience in the eyes of voters and their colleagues may gravitate towards these issues to gain credibility, this will diminish with increased seniority. I present findings that, at the aggregate level, are broadly consistent with past work, and show that women politicians do in fact talk more about traditional social policy issues such as children and family, health, education, women, and welfare than men. However, these aggregate level differences mask significant variations over the course of politicians' careers. I show that, among junior politicians, there are stark gender differences for the social policy areas, however, these gaps decrease markedly as time in parliament increases. Further, these changes are concentrated only among these issue areas, and not among the wider sets of issues politicians raise.

Substantively, I provide an important contribution to existing work on representation and legislative careers by showing that the pressures politicians experience to focus on issues that may be related to their backgrounds and identity diminish over the course of their careers. I show that these changes are not driven by progression into ministerial or committee positions over time, nor are they the result of selection effects in the kinds of politicians who serve longer parliamentary careers, instead they are likely the result of changes in the sets of issues politicians decide to champion throughout their time in office. These findings contribute to myriad existing work that has sought to identify why politicians differ in the issues they enhance and prioritise (Reher, 2022; Sobolewska, McKee and Campbell, 2018), by emphasising the potential importance of experience and incentives for shaping politicians' decision-making.

These findings also have important implications for the substantive representation of women's voices and interests in politics. I find that while women politicians diversify the sets of issues they focus on the longer they are in office away from traditional "feminine" issues, they continue to represent women's perspectives throughout their time in office. Women talk more about women overall, do so in debates on all issue areas, and this difference does not decline with increased seniority. Further, as women politicians become more senior, they champion a wider variety of issues, but also continue to raise women's experiences and perspectives in these issues. This suggests that women's interests are being voiced in debates on a wide variety of issue areas. Notably, some of the largest effects I document are for traditional "masculine" policy areas such as crime and policing, which may suggest that when women participate on debates in these areas, they highlight issues that disproportionately affect women such as failures in adequately policing violence against women.

Methodologically, I combine traditional quantitative dictionary approaches with newer innova-

tions in word-embedding models to create measures of several issues in contemporary British political debate. By focusing on a wide variety of issues, I have examined the potential for gender differences without determining *a priori* the sets of potentially stereotypical issues men and women might raise. The advantage of this approach is that, while I have found clear evidence that women politicians do focus more than men on social policy areas, I do not find that women talk any less about issues such as the economy which prior work has identified as “masculine” (Bäck, Debus and Müller, 2014). While my focus has been on the importance of gender and seniority on the issues politicians raise, the measures developed should prove fruitful for researchers interested in questions such as how politicians’ pre-political experience (Allen, 2013) or constituency preferences (Blumenau, 2021b) may shape the issues they raise.

I have focused on the issues politicians prioritise as measured in the speeches they deliver in parliament. There are several notable advantages to the use of speech data, such as the limited control that party leaders wield over the content of politicians’ speeches, however, a clear disadvantage is that these speeches have little tangible impact on policy outcomes. In the UK, it is challenging to measure the behaviour and motivations of individual politicians through voting behaviour because legislation is commonly proposed centrally by the government and MPs tend to vote in accordance with their party’s position. However, underpinning the argument I present here is that women politicians need navigate an electoral and institutional environment that is more precarious for them than it is for men. Just as this situation is likely the reality for women in the UK, this unfortunately describes many political environments beyond the UK too. The argument I present is a potentially general one that may apply to other legislatures where women are under-represented and where individual contributions to legislative voting behaviour are observable. Understanding the extent to which these behaviours translate into voting behaviour and policy outcomes, or, indeed, to other legislative arenas, is a fruitful avenue for future work, and I encourage future scholars to tackle these questions.

## Chapter 5

# A Double Standard? Gender Bias in Voters' Perceptions of Political Arguments

A version of this chapter is forthcoming as: Hargrave, Lotte (2022) "A Double Standard? Gender Bias in Voters' Perceptions of Political Arguments", *British Journal of Political Science*, 1-19. [https://doi:10.1017/S0007123422000515](https://doi.org/10.1017/S0007123422000515).

### Abstract

Do the styles politicians use influence how voters evaluate them, and does this matter more for women than for men? Politicians regularly use anecdotal arguments, emotional appeals, and aggressive attacks when communicating with voters. However, that women politicians have been branded as "nasty", "inhuman", and "unfeminine" suggests these strategies may come at a price for some. I report on a novel survey experiment assessing whether voters are biased in their perceptions and evaluations of politicians' communication styles. By manipulating politician gender and argument style I assess, first, whether politicians incur backlash when violating gender-based stereotypes, and second, whether differential perceptions of the styles themselves explain this backlash. I find that style usage has important consequences for how voters evaluate politicians, but that this is not gendered. These results have important implications as they suggest that women politicians may not need to conform to stereotype-expected behaviours to receive positive voter evaluations.

## 5.1 Introduction

Discussion and debate lie at the heart of democratic politics. When communicating with the public, politicians make regular use of a wide range of argumentation strategies such as empathetic appeals, emotional arguments, and aggressive attacks. What politicians say, and importantly, how they say it, matters a great deal. While politicians expressing empathy, emotion, and aggression is unremarkable, certain politicians may face harsher penalties than others when they express, or do not express, these styles. Former UK Prime Minister Theresa May was known for a style that was characterised as unemotional, unempathetic, and robotic. May's style received significant press commentary with some asking "what is wrong" with her, others branding her as "inhuman and uncaring", or suggesting her "inability to show emotion to the public proves that she isn't fit to be Prime Minister". In the eyes of the British public too, May was seen as having a "cold personality". Eventually, this led to her infamous characterisation in the press and across social media as "the Maybot". Yet, when her prime ministerial resignation speech was uncharacteristically marked with emotion and tears, it prompted some to bid farewell to a leader who was "almost human after all".

Theresa May's style is one that goes against what is stereotypically expected of women in politics (Eagly and Karau, 2002), and she has been branded as inhuman, cold, and robotic as a result (Harmer, 2021). While this anecdote illustrates only one example of how the styles of women are perceived and evaluated, there is ample evidence that women politicians more broadly incur penalties from voters when they do not conform to stereotype-congruent behaviours (Bauer, 2015a; Boussalis et al., 2021; Cassese and Holman, 2018). Men politicians, by contrast, have been shown to be successful both when they conform to stereotype expected behaviours – such as being ambitious (Saha and Weeks, 2022) – and when they subvert stereotypical expectations – such as being emotional or communal (Gleason, 2020; Okimoto and Brescoll, 2010). Indeed, past work has shown there to be clear evidence of asymmetric standards in how men and women politicians are evaluated.

In this chapter, I assess whether voters are biased in their perceptions and evaluations of the ways in which politicians communicate, and, consequently, whether voters' assessments of elite political communication are gendered. While there is a robust literature on the influence of gender stereotypes on voter *evaluations* in politics more broadly, so far scholars have neglected to unpack whether voters differentially *perceive* politicians' behaviour based on gender alone. This work makes an important contribution by being the first to address this question, and to assess whether differential perception of styles themselves may serve as a potential mechanism through which voters' gendered evaluations



of politicians might become manifest.

To address these questions, I focus on styles for which the gender stereotypes literature has outlined clear expectations about men and women's behaviour (Eagly and Wood, 2012; Schneider and Bos, 2019). Work to date examining how voters evaluate stereotype-(in)congruent behaviour has focused on isolated behaviours (e.g., Cassese and Holman, 2018; Krupnikov and Bauer, 2014). By contrast, while I analyse the effect of each style separately, I make progress by focusing on how voters evaluate politicians' use of a wider range of styles that are consistent with both feminine communal stereotypes and masculine agentic stereotypes (Eagly and Karau, 2002). By doing so, I assess whether there are asymmetric standards in the degree to which men *and* women politicians are punished for stereotype-incongruent behaviours. I focus on the use of emotion, aggression, and evidence (both drawing on *statistical* and *anecdotal* evidence). In a novel survey experiment, I present UK voters with speeches where the argument style and the gender of the politician delivering the argument are varied. Through these manipulations I assess, first, whether politicians incur negative evaluations from voters when they deploy styles that are gender stereotype-incongruent, and second, whether voters' differential perception of styles themselves might explain the backlash politicians receive.

I report four main findings. First, I find that the styles politicians use has important consequences for how voters evaluate them. Politicians are liked more when they are emotional or draw on anecdotes, however, they are regarded as more competent when they are unemotional and make use of statistical evidence. Second, contrary to expectations from the gender stereotypes literature, I find no evidence that these evaluations are gendered. That is, while all politicians are evaluated as less likeable when they are aggressive, there is no evidence that women in particular incur negative evaluations for this behaviour. Third, while there is clear evidence that voters can identify the styles politicians use, I also find no evidence that voters' perceptions of the styles themselves are gendered. Voters do not perceive unemotional arguments by women as any less emotional than unemotional arguments by men, nor do they perceive aggressive arguments by women to be any more aggressive than equally aggressive arguments by men. Fourth, I find some evidence that these perceptions and evaluations differ by voter gender. If anything, women voters reward women politicians for stereotype-congruent behaviour: they give a larger likeability and competence reward to women politicians who are emotional and perceive arguments by women politicians to be more emotional than equivalent arguments delivered by men politicians.

The main finding I document, therefore, is that there is little evidence of gender bias in the forum I study. These findings do not, however, imply that there is no gender bias at play in politics. I study voter

evaluations of the likeability and competence of politicians as gendered and non-gendered accounts of leadership alike suggest they are some of the key qualities expected of politicians (Clarke et al., 2018; Huddy and Terkildsen, 1993a). However, even if voters equally evaluate men and women for the styles they use, I do not assess how likeability and competence may have downstream consequences for voting. Previous US-based work has suggested that competence evaluations may be more important in voters' evaluations of women (Ditonto, 2017), and therefore women may *need* to be perceived as more competent than men to get elected. The media may also play an important role in the *framing* of women's behaviour. If the media frame women's behaviour in a more negative light than men's, this may in turn feed into how voters evaluate politicians even if voters' direct judgements are not themselves gendered. Finally, my results suggest that voters deem politicians to be less competent when they express styles traditionally associated with feminine "communal" stereotypes (Schneider and Bos, 2019). Therefore, if the distribution of actual style usage is different for men and women then this may still lead to bias. If it is true that women happen to make greater use of communal styles – such as emotion or anecdotes – than men, then they may incur negative competence evaluations simply because of the styles they use. Evidence from a variety of contexts has shown that women politicians are more emotional (Dietrich, Hayes and O'Brien, 2019) and draw more on anecdotes (Hargrave and Langengen, 2021), although the results presented in chapter 3 shows that women have made decreasing use of these styles over time in the UK. In short, identifying the mechanism through which biases emerge in the judgement of politicians' behaviour is challenging, and I make an important contribution by shedding light on one crucial aspect of this complex process: whether voters' perceptions of and attitudes towards elite political communication are gendered.

A commonly held perception is that men's leadership styles are preferred for political office than women's (Fox and Oxley, 2003), and that when women try and adopt these styles, they violate feminine stereotypes, and will lose out. However, the findings I present here suggest the reality for women may be more positive than this, as the expectation that they must avoid these styles for fear of negative backlash from voters is somewhat misguided. Recent work has shown that discrimination can occur when individuals perceive that others are likely to discriminate and that voters tend to overestimate the degree to which *others* are biased (Bateson, 2020). Documenting that voters may not be as biased towards stereotype-incongruent women is important, then, as it may potentially help to further reduce voter bias. While of course women may be sanctioned for stereotype-incongruent behaviour from other sources, the findings I present here suggest that voters do not punish women in the way common theories of gender stereotyping may have predicted.

## 5.2 Gender, stereotypes, and voter backlash

Why might voters hold different expectations for how men and women politicians behave? Gender role theory suggests that stereotypes concerning the typical behaviours of men and women lead to strong expectations for how individuals of different genders will and should behave. The social roots of such gender roles are thought to emerge from a type of statistical profiling: to the extent that, for diverse historical reasons, women and men display different behaviours on average, people will internalise these patterns and the “corresponding attributes become stereotypic of women [men] and part of the female [male] gender role” (Eagly and Wood, 2012, 11). These behaviours are not always precisely defined; however, they are broadly divided into women’s communal behaviours which are associated with emotionality, positivity, warmth, human interest, and the wider idea of caring for others, and men’s agentic behaviours which are associated with aggression, logic, and leadership (Schneider and Bos, 2019).

Over time, the expectation that men and women should behave in ways traditionally associated with their gender is reinforced through a process of socialisation. Further, descriptive stereotypes, which purport to describe what group members *are* like, eventually lead to prescriptive stereotypes, which purport to describe what group members *should* be like (Gill, 2004). The idea that women are, say, gentle and emotional eventually leads to the expectation that women should be gentle and emotional. Behaviour that is inconsistent with these stereotypes may then be punished via social sanctions imposed by others. The upshot is that gender roles become self-reinforcing, as societal beliefs about the typical behavioural differences lead to ever-more entrenched patterns of gendered behaviour.

Men and women politicians may therefore be evaluated differently by voters based on their behaviour. I focus on four *styles* for which the gender stereotypes literature outlines clear expectations for men and women’s usage. The first communal style is **emotion**. Women are expected to be more emotional (Dietrich, Hayes and O’Brien, 2019; Huddy and Terkildsen, 1993a), and more positive specifically (Yu, 2013). The second communal style is **anecdotal evidence**. Women are thought to make greater use of anecdotes, which includes referencing personal experience, the experiences of others, analogies, or stories (Blankenship and Robson, 1995; Childs, 2004b; Hargrave and Langengen, 2021). The first agentic style is the use of **aggression**. Men are thought to be more aggressive (Grey, 2002), whereas women are thought to avoid this behaviour for fear that it is “negatively perceived by the electorate” (Childs, 2004b, 190). The second agentic style is the use of **statistical evidence** (Childs, 2004b; Mattei, 1998), which is linked to the idea that men are thought to be more “analytical, organised and

impersonal” (Jamieson, 1995, 76). Table 5.1 includes a definition of each of the styles of interest, which are drawn from the literature described here.

Table 5.1: Style definitions

Style	Definition
Emotion	Positive emotional language, which includes expressing empathy, praise, celebration, congratulations, hope, and joy
Aggression	Language that relates to conflict, political point-scoring, criticisms, or insults
Statistical evidence	Use of numbers, statistics, numeric quantifiers, figures, and empirical evidence as the basis for an MP’s argument
Anecdotal evidence	Use of personal examples or experiences, stories of other people, constituency stories, or illustrative examples as the basis of an MP’s argument

How might the conformity to or violation of stereotype-congruent behaviours affect voter evaluations of politicians? Previous work has shown that women politicians are subject to pressures to conform to stereotypes. When women express behaviours which are *consistent* with stereotypes of communality, they tend to be rewarded. In recent observational work, Boussalis et al. (2021) focus on the behaviour of Angela Merkel and find that she is rewarded by voters for displays of happiness but punished for expressions of anger. Similarly, work by Gleason (2020) shows that women Supreme Court attorneys are most successful when they use extensive emotional appeals. By contrast, when women express behaviours which are *inconsistent* with stereotypes of communality, they can incur backlash for it. For instance, Cassese and Holman (2018) find that women candidates in particular are vulnerable to attacks from voters when they are expressly negative in campaigns. How might these evaluations translate to men? Voters have been shown to be less sensitive to men expressing styles which are incongruent with masculine “agentic” stereotypes (Gleason, 2020), and indeed studies suggest that men are successful when they both follow *and* subvert gendered expectations (Okimoto and Brescoll, 2010).

I seek to identify whether men and women politicians’ use of stereotype-(in)congruent styles leads to differential evaluations from voters. Previous work has focused on a wide variety of traits that might be influenced by conformity to stereotypes (e.g., Brooks, 2011; de Geus et al., 2021; Saha and Weeks, 2022), which can be divided along two lines. First, the extent to which politicians’ conformity to stereotypes leads to voters feeling more *warmly* towards them as individuals. Women are expected to be communal, which includes being kind, caring, and compassionate (Alexander and Andersen, 1993). Therefore, when they behave in ways that are instead consistent with agentic stereotypes, such as

being aggressive or ambitious, they fall short of expectations about the communal qualities deemed appropriate for women (Schneider and Bos, 2019). Women who behave in an overly agentic manner may be negatively evaluated. Studies assessing voter evaluations of role-incongruent women confirm that women are perceived less favourably by voters (Boussalis et al., 2021), whereas role-congruent women are regarded more favourably and rewarded via increased likeability evaluations (Bauer, 2017).

Second, conformity to stereotype-congruent behaviours may also influence the extent to which voters judge politicians' ability to perform their jobs to a high standard. Politicians are expected to be competent and, as the traditional occupants of political leadership positions, men's competence in these roles is assumed in a way that women's is not (Schneider and Bos, 2019). Recent experimental work has shown that women *overall* may actually have a small electoral advantage relative to men (Schwarz and Coppock, 2022), that voters may be no more critical of women politicians' past performance in office than men's (de Geus et al., 2021; Hargrave and Smith, 2022), and that stereotypes about women's competence may have equalised in recent years (Donnelly et al., 2016). Nonetheless, a large literature has shown that women candidates have historically been evaluated as less competent (Alexander and Andersen, 1993; Huddy and Terkildsen, 1993a; Koenig et al., 2011), and that voters seek out more information about the qualifications of women than men (Ditonto, 2017; Ditonto, Hamilton and Redlawsk, 2014). Further, in the campaign context, recent experimental work has shown women candidates require stronger qualifications than men (Bauer, 2020a). Women may, therefore, need to do more to prove their competence in their political roles. Consequently, I focus on both likeability and competence, which gendered and non-gendered accounts of leadership alike identify as some of the key qualities expected of politicians (Clarke et al., 2018; Huddy and Terkildsen, 1993a). For both, I expect that women politicians will be rewarded by voters when they express styles which are congruent with feminine communal stereotypes and be punished when they express styles which are congruent with masculine agentic stereotypes.

This study contributes to the literature discussed above in three key ways. First, work to date examining how voters evaluate stereotype-congruent or incongruent behaviour has focused on isolated behaviours, such as tears (Brooks, 2011), negative attacks (Krupnikov and Bauer, 2014), or ambition (Okimoto and Brescoll, 2010).<sup>1</sup> By contrast, I expand upon this to focus on how voters evaluate politicians' use of a wider range of styles that are consistent with both feminine communal and masculine agentic stereotypes. In doing so, I assess whether there are asymmetric standards in the degree to

---

<sup>1</sup>Though work on campaign strategy has examined how the use of stereotype-(in)congruent issues and traits may have different successes (Schneider, 2014a,b).

which men *and* women politicians are punished for stereotype-incongruent behaviours.

Second, to date, work on voter backlash to stereotype-incongruent behaviour has focused almost exclusively on the US case (Alexander and Andersen, 1993; Bauer, 2015a, 2020a; Bauer, Kalmoe and Russell, 2022; Bauer and Taylor, 2022; Cassese and Holman, 2018; Ditonto, Hamilton and Redlawsk, 2014; Ditonto, 2017; Huddy and Terkildsen, 1993a; Krupnikov and Bauer, 2014; Mo, 2015). So far, European work has studied gender bias in voting behaviour (Campbell and Cowley, 2014; Campbell et al., 2019a,b; Vivyan and Wagner, 2015). However, there is growing attention to the degree to which UK voters hold politicians to stereotypical standards (Eggers, Vivyan and Wagner, 2017; Hargrave and Smith, 2022; Saha and Weeks, 2022; Smith, 2019), which has near consistently uncovered that voters do not seem to stereotype politicians. Given that gender norms and stereotypes are said to be dynamic concepts that shift over time and context (Diekman and Eagly, 2000; Eagly et al., 2020), I contribute to existing work by testing new aspects of theory on gender stereotypes drawing on data from the UK case.

Third, while there is a robust literature on the influence of gender stereotypes on voter evaluations in politics more broadly, so far scholars have neglected to unpack whether voters differentially perceive the behaviour of men and women, and whether differential perception of behaviour itself may in turn be responsible for how voters evaluate politicians. This study is the first to assess whether differential perceptions of styles themselves may serve as a potential mechanism through which voters' gendered evaluations of politicians might become manifest.

### *Differential perceptions of styles*

What may explain why voters differentially penalise men and women politicians for expressing styles that violate stereotypes? One mechanism through which these evaluations might become manifest is that – because of stereotypical expectations – voters simply perceive differently the styles that men and women use. That is, even in the absence of any *objective* differences in, say, the emotionality of a speech, voters will “hear” that speech as more or less emotional if delivered by a woman or a man. This differential perception of style may then, in turn, lead to voters' differential evaluation of the likeability or competence of men and women.

Aggression is a style that is congruent with masculine stereotypes. Therefore, when voters see a man delivering an aggressive speech, this may only serve to confirm pre-existing stereotypes. Stereotypical expectations dictate, however, that women are *not* expected to be aggressive. Voters may therefore pay more attention to women when they behave aggressively because it violates stereotypes (Brescoll and Uhlmann, 2008). Because voters see a woman as particularly aggressive this may,

in turn, affect their likeability or competence evaluations of that woman. The experiences of former US presidential candidate Hillary Clinton – who was described as “too angry, aggressive, and unfeminine” (Brooks, 2013, 1) – may help to further explain this dynamic. Clinton may have been perceived as more aggressive than her men counterparts even when she may not actually have behaved in such a way because aggression is congruent with masculine stereotypes. Therefore, because Clinton’s use of aggression is particularly noticeable, this may feed into her being evaluated more negatively. Work on Canadian elections supports this perspective and finds that women leaders’ speeches are reported by the media as more negative and aggressive (Gidengil and Everitt, 2003).

Similarly, emotion is a style that is congruent with feminine communal stereotypes. Therefore, when women are not emotional, this may violate expectations about the supposed emotional sensitivity of women and be of particular note to voters. Because women are therefore seen as particularly unemotional, they, relative to men, may be ascribed more negative evaluations and be seen as particularly “cold and unlikeable”.

I also expect that there will be differences in the degree to which voters perceive men and women politicians’ use of anecdotes and statistics as evidence-based. While existing work has presented mixed conclusions about the persuasiveness of these evidence types, it is broadly concluded that statistical arguments help ensure argument credibility (see Hornikx, 2018, for a discussion). As outlined above, studies have highlighted that men’s competence in office is often assumed, while women’s is not. Therefore, while I expect that arguments by *all* politicians will be perceived as more evidence-based when they contain statistical arguments over anecdotal arguments, arguments by women politicians *in particular* will be perceived as more evidence-based when they include statistical evidence compared to anecdotal evidence. However, because of the assumed competence of men, the difference between anecdotal and statistical arguments delivered by men politicians will be smaller.

In the experiment, described next, I assess whether voters perceive men’s and women’s styles differently even in circumstances where there are no differences, and whether differential perception of styles may work as a potential mechanism for explaining variation in the likeability and competence evaluations that politicians receive.

### 5.3 Experimental design

I test my expectations with a vignette survey experiment. In the experiment, respondents were tasked with reading a short argument delivered by a fictitious MP. Respondents were randomly assigned to

different treatment conditions where four attributes were randomised. First, the **style** of interest: *emotion, aggression, or evidence*. Second, the **treatment status** of the style. For emotion, this is the *control* condition which is a neutral speech not representative of emotion and the *treatment* condition which is highly representative of emotion. For aggression, this is the *control* condition which is a neutral speech not representative of aggression and the *treatment* condition which is highly representative of aggression. For evidence, the *control* condition is statistical evidence and the *treatment* condition is anecdotal evidence.

Third, the **policy area**: *housing, health, or transport*. While I am not interested in making inferences about specific policies, I include several as there may be a concern that any effects uncovered for, say, women and emotionality on health may not translate to another area such as transport. Work investigating the relative persuasiveness of different rhetorical techniques found there to be a large degree of heterogeneity in the persuasiveness of different rhetorical elements across policies (Blumenau and Lauderdale, 2022b). Therefore, by including a range of policies, I can average over them to address my central questions. Further, prior US-based work has highlighted that voters may make inferences about a politician's party based on the policy in question as the Democrats have "ownership" of particular issues such as education, and the Republicans over others such as defence (Petrocik, Benoit and Hansen, 2003). We therefore may be concerned that voters will evaluate politicians based on their perceptions of parties. In the UK, while certain issues are arguably associated with a particular party – for instance, welfare and the Labour Party (O'Grady, 2022) – I select housing, health, and transport as issues that are central in British politics but not "owned" by either party. Further, the treatments were written to reflect both Conservative and Labour priorities on the issue areas. For instance, the housing treatments include reference to areas that Labour champions, such as the gaps between wages and house prices and **protecting vulnerable renters**, and issues that the Conservatives have greater authority on, such as **increasing homeownership**. By doing so, I aim to minimise the likelihood that voters make inferences about a politician's party.

Fourth, the **gender** of the MP: *man or woman*.<sup>2</sup> Gender is delivered through fictional Anglo-Saxon names that represent the "typical" MP in the House of Commons. Previous experimental work has shown that changing name alone is a sufficient cue to move voters' attitudes towards a politician's gender (see Campbell and Cowley, 2014, for a discussion). In total, there are 3 x 2 x 3 x 2 (style x

<sup>2</sup>An alternative version of this design could have included a "neutral" control condition, where the gender of the MP was not stated. However, MPs "in the wild" in the UK House of Commons are majority male, and it therefore seems likely that respondents may assume the "neutral" MP is a man. Results from an experiment pre-testing survey (see the appendix for further details) suggest this is indeed the case. In the task, respondents were assigned speeches by MPs without the gender of the MP stated and asked "If you had to guess, would you say this MP is a man or a woman?". Out of all respondents, 56% guessed the MP was a man and 44% guessed the MP was a woman.



treatment status x policy area x gender) = 36 unique treatment conditions which serve as the basis of this experimental design. The full text of all treatments are presented in table C.3.

To ensure that the speeches are similar in spirit to the kinds of arguments politicians use in the UK parliamentary context, the texts of the speeches were informed by searching for debates on similar policy areas recorded in Hansard, the official report of all parliamentary debates. When writing the speeches, the basic structure remained the same within a policy area, however, I added words and sentences deemed representative of the styles. The word choice was informed by the words considered most representative of the styles measured in chapter 3. For instance, words such as “disgraceful”, “deplorable”, and “incompetent” were detected as the most aggressive according to the word-embedding based measure, and are therefore included in the treatment texts. Further, in the UK context, a very common approach for politicians using anecdotes in their arguments is to refer to the experiences of their constituents (Atkins and Finlayson, 2013), and therefore for each of the anecdotal arguments I refer to the same fictional constituency couple that the MP had spoken with. The aggression treatments all begin with the same opening sentences: “I have to start by saying I utterly disagree with what you’ve just said. You clearly lack any understanding of how serious this is, and, frankly, you show absolutely no care for the people we represent.” Table 5.2 shows an example for each style, where the “treatment” of the style is indicated in bold and square brackets. Overall, the aim was to ensure that the treatment was consistent across policy areas.

To ensure the treatments fulfilled these aims, a pre-test survey was fielded by Prolific to 1,499 members of their UK online panel in August 2021. The purpose was to ensure that the treatment texts were considered representative of the styles, and that the control texts were not. Overall, the results confirm that the treatments are perceived as representative of the styles. For instance, 72.5% of respondents assigned to the emotion treatment conditions (strongly) agreed that the treatments were emotional. Similarly, 67.7% of respondents assigned the aggression treatment conditions (strongly) agreed that the treatments were aggressive. While most respondents perceived that both the statistics and anecdote speeches were evidence-based, a larger percentage of respondents perceived the statistical speeches as evidence-based (80.8% and 56.8% respectively). Further, the pre-testing results provide good evidence that the controls were not perceived to be particularly representative of the styles. Only 18.3% of respondents assigned to the emotion control conditions (strongly) agreed that the treatments were emotional. Additionally, 24.4% of the respondents assigned to the aggression control conditions (strongly) agreed that the treatments were aggressive. Full details of the exercise can be found in appendix section C.1.

Table 5.2: Vignette examples

Attributes	Vignette text
Emotion, treatment, transport, woman	<p><b>Lucy Craddock, MP:</b> “Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. <b>[Imagine stepping outside on a bright, beautiful morning and hearing not engines revving nor choking on polluted air but feeling that simple joy of hearing the birds sing and that rush of fresh air into your lungs. Doesn’t this sound amazing? This could be our future, and I so hope that this doesn’t have to be a distant dream.]</b> We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer. <b>[We can fill our lives with the simple pleasures of bird song in our ears, fresh air in our lungs, and blue skies ahead.]”</b></p>
Aggression, treatment, health, man	<p><b>Jack Richards, MP:</b> <b>“I have to start by saying I utterly disagree with what you’ve just said. You clearly lack any understanding of how serious this is, and, frankly, you show absolutely no care for the people we represent.]</b> The way we understand diseases and how to treat them has grown over the years. There have been advances in treatments that might help save the lives of people in this country and we should be committed to using them. <b>[But often nothing is done because it is simply deemed not worth the money. I am utterly revolted by the idea that people are not getting treated merely because of some appalling cost-benefit calculation. This is disgusting, deplorable, and inhuman.]</b> It is important that how we fund and pay for drugs works for everybody. <b>[Things must change, and anyone who opposes this cannot claim they care about the well-being of our people.]”</b></p>
Evidence, statistics, housing, man	<p><b>Adam Jones, MP:</b> “Our housing market has not been as it should for years. It is the job of all of us to increase the supply of housing. But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. <b>[Those in their mid-30s to mid-40s are three times more likely to be renters than 20 years ago. People in the early 1990s could expect to pay 3.5 times their annual earnings on buying a home, but this has risen to 7.8 today. In 2019, the average property sold for £235,300, meanwhile the average pay came in at £29,000.]</b> Our housing market does not work for many people. We need different strategies to help to increase supply and make homes more affordable. There is a need for new policy to help homeownership become realistic for young people all around the country.”</p>
Evidence, anecdote, transport, woman	<p><b>Charlotte Richards, MP:</b> “Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. The market for electric vehicles is small yet growing. We can be industry leaders in how we produce and use electric vehicles. <b>[I spoke recently to Eleanor and Michael, a young couple in my constituency who live in a flat in a high-rise building. They told me while they want to make the swap to an electric vehicle, it is just not practical as they do not have easy access to a charging point.]</b> We should widen accessibility in the use of electric vehicles to make them more practical for people, <b>[just like Eleanor and Michael],</b> who live living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer.”</p>

Attributes	Vignette text
Emotion/aggression, control, housing, man	<p><b>Peter Jones, MP:</b> <i>“The way we understand diseases and how to treat them has grown over the years. There have been advances in treatments that might help save the lives of people in this country and we should be committed to using them. But in some cases, patients are left without access to the drugs they need because of the costs of treatment. Patients being able to access the treatments they require is one of the principles of providing universal health care to everyone who needs it. I think some people find it challenging to accept that questions of money should enter decisions about health care, but this is the situation that many hospitals, doctors, and patients find themselves in. It is important that how we fund and pay for drugs works for everybody. Change should happen to make sure that both medical services and patients get the treatments and resources they need.”</i></p>

I leverage the use of styles in parliamentary speeches which is only one forum of political speech where we might anticipate gender bias. An alternative forum may have been speeches that politicians deliver in political campaigns, or media descriptions of politicians' behaviour. Presenting voters with an alternative forum of political communication, such as media reporting, would enable me to assess only how gendered *framing* of politicians may influence voter judgements, and not how voters may form gendered judgements when they engage directly with the speeches politicians make. While political campaign speeches may seem the more natural forum through which politicians communicate with voters, previous work has highlighted that the single-member district electoral system in the UK provides MPs with an incentive to cultivate a personal vote (Kam, 2009). Further, work on parliamentary speech in the UK has shown that politicians make strategic use of this forum to appeal to voters (Blumenau, 2021b). Recent work by Osnabrügge, Hobolt and Rodon (2021) has shown UK politicians “play to the gallery” and use emotive rhetoric most in debates that reach the widest public audiences to better appeal to voters. Therefore, in the UK case, parliamentary speech is an appropriate forum to study my key questions of interest.

I focus on how voters judge men and women politicians' use of stereotype-(in)congruent styles when they *read* arguments and am therefore unable to make statements about how such dynamics may differ when *hearing* or *watching* politicians. While a central part of style use is captured in the content of what is said, both tone and body language are, of course, also important stylistic features. In recent work, Boussalis et al. (2021) assess whether voters are biased in how they evaluate the styles of Angela Merkel through utilising a variety of video, audio, and text approaches. Their results – that Merkel is punished for aggressiveness and rewarded for happiness – hold across the different measures of style but are most pronounced for non-verbal communication. By focusing on only written speech, it is possible that I may understate the extent to which voters are biased in their judgements of

politicians. Presenting voters with speeches delivered in an audio or video format by men and women may, however, introduce unwanted confounders into the relationship – such as difficulties in holding tone, pitch, or cadence constant – which would be extremely challenging to account for. Instead, by focusing on only written speech, I can hold constant all features of speech except for the gender of the politician delivering the argument, and therefore identify my key substantive quantity of interest: whether a politician’s gender alone influences voters’ perceptions and evaluations.

A further source of concern may be that, although the MP’s party is not stated, voters may make inferences about party based on their gender. In the US, voters tend to stereotype the Democrats as more “feminine” and the Republicans as more “masculine” (Hayes, 2011; Winter, 2010), and voters use a candidate’s gender to infer candidate ideology (Huddy and Terkildsen, 1993a; Koch, 2000; McDermott, 1997). US-based experimental work has also emphasised the importance of partisanship in the degree to which voters stereotype politicians (Bauer, 2018; Cassese and Holman, 2018; Krupnikov and Bauer, 2014). In the US, therefore, voters have been shown to both infer that women are more liberal than men, and that a candidate’s partisanship determines the extent to which they incur negative evaluations. While I am unaware of work that has assessed this question directly in the UK, several factors suggest this is perhaps less of a concern in my context. First, UK right-wing parties have made increasing efforts to integrate women numerically and to better represent women’s interests (Celis and Childs, 2012; Childs and Celis, 2018). While Labour made earlier efforts to increase women’s representation (Childs, 2004b), the Conservatives have also made a concerted effort in recent years to “feminise” the party by incorporating women into the party hierarchy and policy (Campbell and Childs, 2017; Childs and Webb, 2012). Second, while Labour have proportionally more women legislators, the Conservatives have had two women Prime Ministers. Therefore, unlike in the US, the parties are more equal with respect to the visibility of women. Finally, recent work has also shown that while there are important differences with respect to party and gender stereotypes in the US, parties in the UK are less divided (Saha and Weeks, 2022). Overall, gender equality is a less polarising and party-political issue in the UK, and the British public has greater familiarity with the presence of women in both major parties. While party is a critical factor with respect to gender and evaluations in the US, this is perhaps less of a concern in the UK political environment.

### *Survey*

I use these treatments as the basis of a vignette survey experiment which was fielded by YouGov to the UK online panel from September 9th–10th, 2021. I pre-registered the design, expectations, and

analysis plan.<sup>3</sup> The sample was 1,676 people who are nationally representative of the British public on a range of attitudinal and demographic criteria.

Following an introduction screen describing the task, respondents were presented with an argument delivered by a fictional MP. To encourage respondents to read the full speech, there was a 15 second delay between the presentation of the speech and the first question. For a respondent's first task, a style was sampled from the full set of three styles (emotion; aggression; evidence). For the selected style, a policy area was sampled from the full set of three areas (transport; housing; health). Within the policy area, they were assigned a style treatment status (control; treatment). Finally, they were assigned the gender of the MP delivering the speech (man; woman).

Each respondent completed the task three times in total. For a given respondent, styles and policies were sampled without replacement from round-to-round, such that once a respondent had read a speech representative of a style or policy area, they did not read a second speech representative of the same style or policy area. For example, if a respondent was assigned emotion on the first response, they could be assigned only aggression or evidence on the second response, and the remaining style on their final response. Similarly, if a respondent was assigned transport on the first response, they could only be assigned housing or health on the second, and the remaining policy area on their final response. Style treatment status and MP gender were sampled with replacement, such that a respondent could be assigned three treatment speeches or three speeches by a woman MP.

Per task, each respondent was asked three questions. First, for all styles, the respondent was asked whether they "agree or disagree that the MP seems likeable". Second, for all styles, the respondent was asked whether they "agree or disagree that the MP seems competent". Third, the respondent was asked whether they "agree or disagree that the argument made by the MP is [emotional/aggressive/evidence-based]". Respondents were only asked how emotional, aggressive, or evidence-based they found an argument to be for the style assigned. For instance, if assigned the emotion style, they were asked "do you agree or disagree that the argument made by [MP name] is emotional", but not whether they perceived the argument to be aggressive or evidence-based. An example prompt is shown in figure 5.1.

Given the three observations per respondent and 1,676 respondents, the total number of observations is 5,028 and 1,676 per style.

---

<sup>3</sup>The pre-registration documents (Hargrave, 2021) can be found at <https://osf.io/bfds7>.



---

**Charlotte Craddock, MP:** *"Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Last year, the transport sector accounted for 29.8% of total carbon dioxide emissions. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. In London alone, there are 9,400 premature deaths every year because of poor air quality. Electric vehicles offer one method of reducing emissions as they produce no air pollution. The market for electric vehicles is small yet growing. Last year saw the biggest ever annual increase in electric vehicles, with a growth of 175,000 new vehicles, which was up 66% on 2019. We can be industry leaders in how we produce and use electric vehicles. We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer."*

Do you agree or disagree that the argument made by Charlotte Craddock MP is evidence-based?

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree
- Don't know



Figure 5.1: Example experiment prompt (Evidence, statistics, transport, woman)

## 5.4 Methodology

### *Outcome and explanatory variables*

I have five outcomes of interest, all of which are measured on 5-point Likert scales that range from “strongly disagree”, “disagree”, “neither agree nor disagree”, “agree” to “strongly agree” and include a “don’t know” option. Likert scales were selected to maximise respondent interpretation of each of the scale points (Pasek and Krosnick, 2010). First, the **likeability** evaluation of an MP. Second, the **competence** evaluation of an MP. Third, the **perceived emotion** of an argument. Fourth, the **perceived aggression** of an argument. Fifth, the **perceived evidence** of an argument. I drop all “don’t know” responses in the analysis.

There are two main explanatory variables. First, the **gender** of the MP, which takes the value of 0 for a man and 1 for a woman. Second, the **style treatment**, which describes the treatment group status of the style, and takes the value of 0 for the control conditions (non-emotional style, non-aggressive style, and statistical evidence) and 1 for the treatment conditions (emotional style, aggressive style, anecdotal evidence).

### *Empirical strategy*

Responses from the emotion, aggression, and evidence styles are modelled separately as the expected direction of the gender effect differs depending on the style. For each, I estimate a series of OLS regression models to investigate my key quantities of interest. First, to assess whether style usage affects voters’ evaluations of politicians and perceptions of their arguments, I estimate a series of models for the five outcomes  $Y_{i(j)}$  for an individual  $i$  in a style  $j$  of the following form:

$$Y_{i(j)} = \alpha + \beta_1 TreatmentStyle_j + \epsilon_i \quad (5.1)$$

where  $\alpha$  in each model describes the average perceived emotion, aggression, evidence, likeability, or competence in the control conditions (non-emotional style, non-aggressive style, and statistical evidence), and  $\alpha + \beta_1$  describes the same quantities in the treatment conditions (emotional style, aggressive style, and anecdotal evidence).

Second, to assess whether the effects of style usage on voters’ evaluations of likeability differ by

the gender of the MP they are evaluating, for each style I estimate the following:

$$\begin{aligned} Likeability_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 TreatmentStyle_j + \\ & \beta_3 (WomanMP_j \cdot TreatmentStyle_j) + \gamma X_{i(j)} + \epsilon_i \end{aligned} \quad (5.2)$$

where  $\beta_1$  in each model describes the difference between men and women when using the control styles (non-emotional, non-aggressive, and statistical evidence).  $\beta_2$  describes the difference between using control and treatment styles (emotional style, aggressive style, and anecdotal evidence) among men.  $\beta_3$  is the key quantity of interest, which describes the difference in the effect of using treatment styles compared to control styles for women compared to men. I expect the  $\beta_3$  coefficients in the emotion and evidence models to be positive, as emotion and anecdotes are female stereotype-congruent styles. In aggression, I expect it to be negative, as aggression is incongruent with feminine stereotypes and women should therefore suffer in likeability evaluations compared to men.  $X_i$  is a vector of additional respondent covariates (gender, age, and education).<sup>4</sup>

Third, to assess whether the effects of style usage on voters' evaluations of competence differ by the gender of the MP they are evaluating, for each style I estimate the following:

$$\begin{aligned} Competence_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 TreatmentStyle_j + \\ & \beta_3 (WomanMP_j \cdot TreatmentStyle_j) + \gamma X_{i(j)} + \epsilon_i \end{aligned} \quad (5.3)$$

where  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  describe the same quantities as above. I expect  $\beta_3$  in the model for emotion to be positive, suggesting that when women conform to feminine stereotypes, they receive a greater competence reward than men. For aggression, I expect  $\beta_3$  to be negative, suggesting that women incur negative competence evaluations when they violate feminine stereotypes. For evidence,  $\beta_3$  will be negative as I expect that women are rewarded in competency evaluations when they deliver statistical as opposed to anecdotal arguments.  $X_i$  is a vector of additional covariates (gender, age, education, and political attention).

Finally, to assess whether a politician's gender leads to differences in  $PerceivedStyle_{i(j)}$ , for each

<sup>4</sup>Across the various models, there are six additional covariates. First, the **policy** area of the argument (categorical: transport; housing; health). Second, **respondent gender** (binary: man; woman). Third, **respondent left-right placement** (continuous: 1-7). Fourth, **respondent age** (continuous: age in years). Fifth, **respondent education** (binary: no degree; degree). Sixth, **respondent political attention** (continuous: 0-10). As outlined in the pre-analysis plan, pre-treatment covariates are included as they should, in expectation, increase the precision of the analysis by explaining variation in the outcome variables.



style I estimate the following:

$$\begin{aligned} PerceivedStyle_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 TreatmentStyle_j + \\ & \beta_3(WomanMP_j \cdot TreatmentStyle_j) + \gamma X_{i(j)} + \epsilon_i \end{aligned} \quad (5.4)$$

where  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  describe the same quantities as above.  $X_i$  is a vector of additional covariates (respondent gender, respondent left-right placement, respondent age, and respondent education).

In the empirical strategy described above, I outline numerous statistical tests that will be conducted, and there is risk of the multiple comparisons problem. To assess whether the results are robust, I carry out subsequent analyses with adjusted  $p$ -values which control for the False Discovery Rate using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995).<sup>5</sup> The results reported below are unadjusted, and the adjusted  $p$ -values are reported in table C.10 in appendix section C.5.

## 5.5 Results

### *Unconditional effects*

In figure 5.2, I present the results from the models described above in equation 5.1.<sup>6</sup> There are three findings to note. First, the top panel – which show the estimated difference between treatment and control styles for the style perception outcomes – highlight that the style treatments were successful in shifting perceptions. The aggressive and emotional styles were perceived as significantly more aggressive and emotional than the non-aggressive and non-emotional styles. Further, statistical evidence was perceived as significantly more evidence-based than anecdotal evidence. Voters therefore do perceive the treatment arguments as more representative of the styles than the control arguments.

Second, the styles that politicians use influences voters' evaluations of their likeability: aggressive politicians are significantly less likeable than non-aggressive politicians, and politicians using anecdotes are significantly more likeable than politicians using statistical evidence. While the point estimate suggests emotional politicians are more likeable than non-emotional politicians, the effect is not statistically significant.

<sup>5</sup>There are other approaches to addressing the multiple comparisons correction. A common solution is the Bonferroni correction. However, the Bonferroni correction assumes independence between tests conducted which is inappropriate in the case of this experimental design, where any form of gender-bias driving perceptions of women's use of aggression is also likely to inform how respondents perceive and evaluate women's use of statistical evidence (Gelman, Hill and Yajima, 2012).

<sup>6</sup>All results from the analyses into the conditional relationship between style treatment status and MP gender and the various outcomes reported in later subsections were pre-registered as documented in Hargrave (2021). However, the analyses into the unconditional relationship between style treatment status and the outcomes presented in this subsection were not pre-registered.

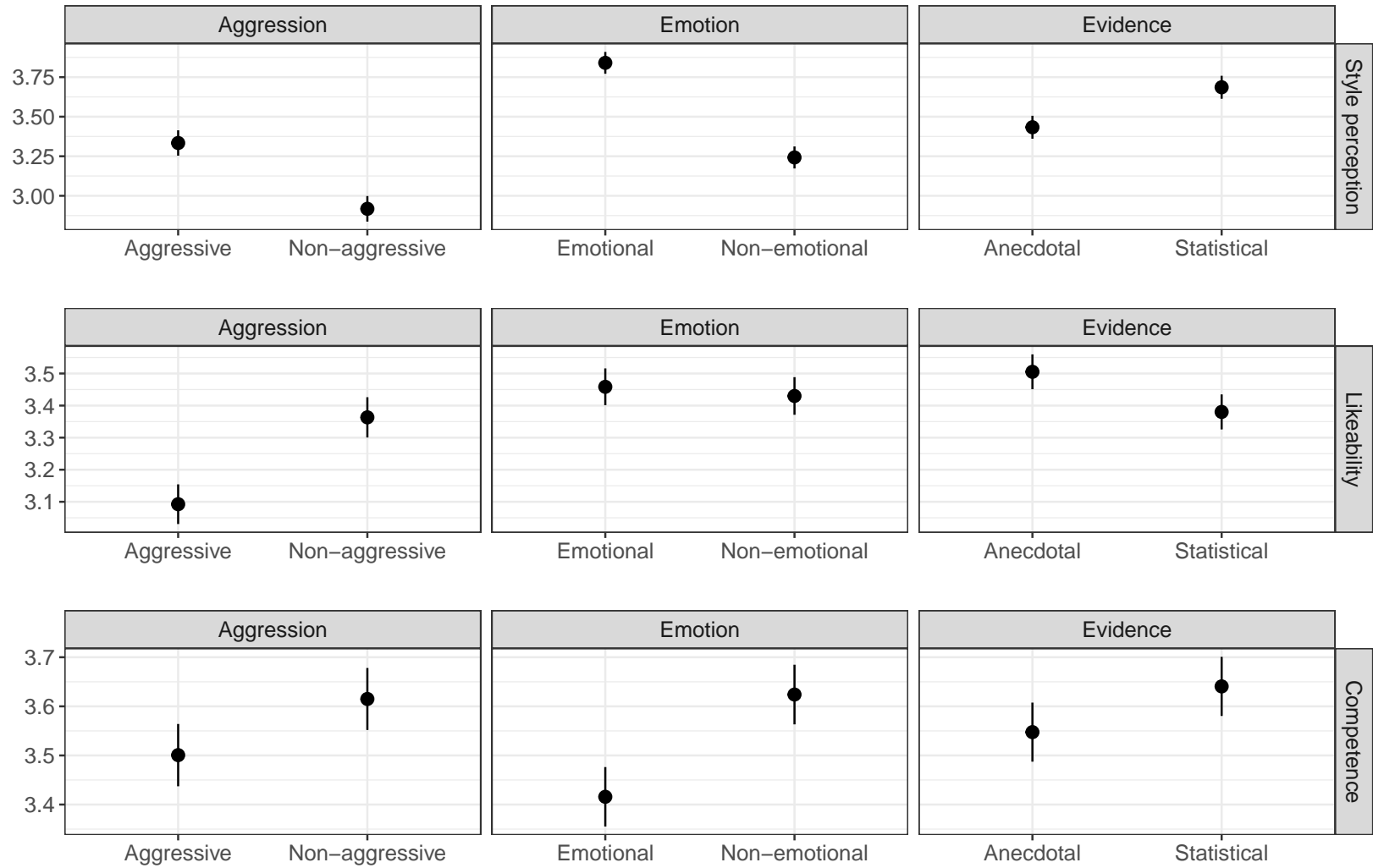


Figure 5.2: Unconditional relationship between control and treatment style usage and style perceptions, MP likeability, and MP competence evaluations

Third, in the bottom row, I document how style usage influences voters' competence evaluations. Politicians using non-aggressive and non-emotional language are evaluated as more competent than those using aggressive or emotional language. Further, politicians are evaluated as more competent if they use statistical evidence as opposed to anecdotal evidence. Therefore, the styles that politicians use influences how voters evaluate them. However, it seems that certain styles lead to trade-offs in evaluations: for instance, the use of statistical evidence leads to politicians being perceived as *less* likeable but *more* competent. Yet, aggression seems not to afford politicians positive evaluations with respect to either competence or likeability. Prior work has shown that aggressive behaviour is unpopular with the public (Dahlerup, 2017; Hansard Society, 2014), that voters dislike uncivil candidates (Bauer, Kalmoe and Russell, 2022), and is relatively unpersuasive as an argumentative technique (Blumenau and Lauderdale, 2022b). My results therefore lend further support to these conclusions about the unpopularity of aggressive language usage by politicians.

#### *Conditional effects by MP gender*

Does style usage matter more for women than men? Figure 5.3 shows the results for the likeability outcomes in the top row of plots.<sup>7</sup> I expected that women would be evaluated as more likeable when they express styles which are congruent with female stereotypes and, conversely, suffer in likeability evaluations when they violate stereotypes. The top-left panel shows the results for aggression. Here, both men and women politicians are punished for being aggressive. The effect is larger for women than for men, but this difference is not statistically significant. The top-middle panel shows the results for emotion. Here, while the direction of the effect suggests that emotional arguments improve likeability amongst women but not men, the effect is not statistically significant. Further, as expected, men are *not* penalised when they use stereotype-incongruent styles (non-emotional). The top-right panel shows the results for evidence. Here, I again do not see that women are penalised when they express styles which are incongruent with female stereotypes (statistical evidence) compared to styles which are congruent with female stereotypes (anecdotal evidence).

Therefore, across the styles, there is no evidence that women politicians are disproportionately penalised in likeability assessments for using styles which are stereotype-incongruent (non-emotional, aggressive, and statistical evidence). As figure 5.2 shows, there is, however, good evidence that the styles politicians use does affect voters' likeability assessments. When politicians use styles consistent with communal stereotypes they are perceived as *more* likeable than styles consistent with agentic

<sup>7</sup>See tables C.7–C.9 in the appendix for full model outputs.

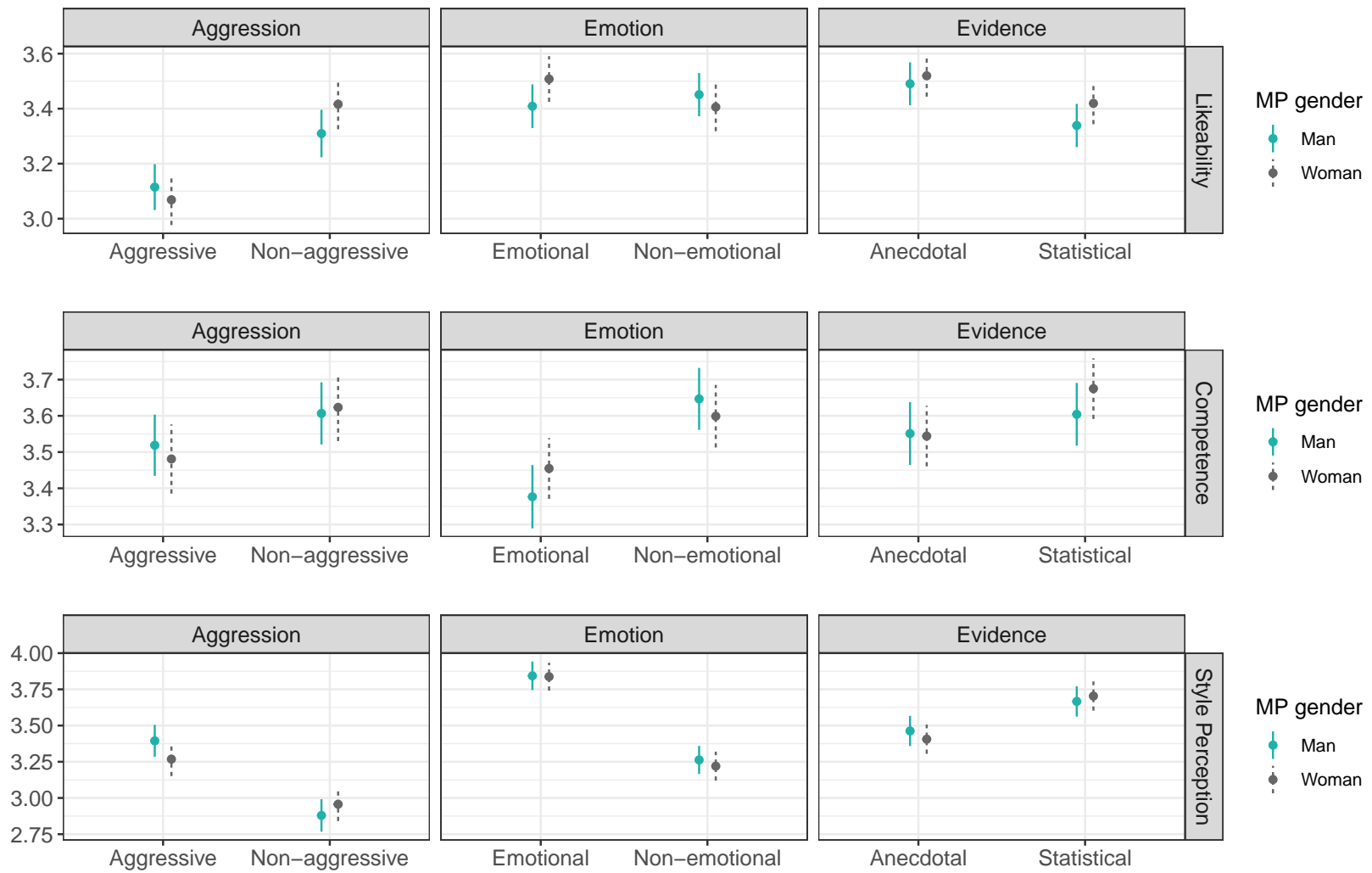


Figure 5.3: Conditional relationship between MP gender, style treatment group, style perceptions, MP likeability, and MP competence evaluations

stereotypes.<sup>8</sup> That voters find politicians to be more likeable when they express styles consistent with the concept of communality may not be a surprising finding, given that communal stereotypes are associated with being warm, kind, emotional, and people-oriented (Schneider and Bos, 2019).

The middle row of plots in figure 5.3 show the results for the competence outcomes. In the middle-left panel, I show the results for aggression. There is no evidence that either men or women MPs are perceived as more competent when they are aggressive than when they are not. For emotion, there is no evidence that women are evaluated as more competent when they express styles which are congruent with female stereotypes (emotional style). However, the use of female stereotype-congruent styles results in men being evaluated as less competent than female stereotype-incongruent styles. For evidence, which is displayed in the middle-right panel, I again see that neither men nor women MPs are evaluated as more competent when they use either anecdotal or statistical evidence.

Therefore, as with likeability, I find no evidence that women MPs in particular are penalised in competence evaluations when they express female stereotype-incongruent styles. As with likeability, voters' competency evaluations of politicians are affected by the styles they use. When politicians express styles consistent with communal stereotypes they are perceived as *less* competent than styles consistent with agentic stereotypes.<sup>9</sup> That voters find politicians to be more competent when they express styles consistent with agentic stereotypes again seems to be an intuitive finding given the compatibility between agentic stereotypes and leadership stereotypes (Bauer, 2017). The findings for emotion and aggression are consistent with work by Brooks (2011), who finds no double standard in the extent to which voters penalise men and women politicians for their expressions of anger and tears.

In the appendix in section C.6, I also estimate models pooling all styles together where the outcomes are the likeability and competence evaluations, the explanatory variables are a categorical variable for each style level, MP gender, the interaction between the two, and controls for policy areas.<sup>10</sup>

<sup>8</sup>When estimating a model pooling all styles together where the outcome is the likeability evaluation and the explanatory variable is whether a style is female stereotype-congruent (emotional style, non-aggressive style, and anecdotal evidence) or not (non-emotional style, aggressive style, and statistical evidence), I find that politicians are perceived as *more* likeable when they express female stereotype-congruent styles. See appendix section C.6 for the full model output.

<sup>9</sup>When estimating a model pooling all styles together where the outcome is the competence evaluation and the explanatory variable is whether a style is female stereotype-congruent, I find that politicians are perceived as *less* competent when they express female stereotype-congruent styles. See the appendix for the full model output. See appendix section C.6 for the full model output.

<sup>10</sup>In this analysis, I pool all styles together and compare each style to the pure control texts for each policy area. To do so, I first create a categorical variable for the **style type** of an argument, which takes the values of control (0), emotional style (1), aggressive style (2), statistical evidence (3), and anecdotal evidence (4). With this variable in hand, I estimate the following:

$$\begin{aligned}
 Likeability_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 Emotion_j + \beta_3 Aggression_j + \\
 & \beta_4 Statistics_j + \beta_5 Anecdote_j + \beta_6 (WomanMP \cdot Emotion) + \\
 & \beta_7 (WomanMP \cdot Aggression) + \beta_8 (WomanMP \cdot Statistics) + \\
 & \beta_9 (WomanMP \cdot Anecdote) + \gamma X_{i(j)} + \epsilon_i
 \end{aligned}
 \tag{5.5}$$

This analysis enables me to compare the effect of each of the styles back to the control arguments for each of the policy areas. While there are some interesting findings with respect to the policy areas – for instance, voters find politicians delivering arguments on the housing and health policy areas to be both more likeable and competence than on the transport policy area – the upshot of this analysis is consistent with the main results presented here. That is, politicians’ style usage influences voters’ evaluations of their likeability and competence, but these evaluations are not gendered. I present further details and a full interpretation of this analysis in the appendix in section C.6.

The bottom row in figure 5.3 shows the results from the style perception models. In the left-hand plot, voters do perceive politicians as more aggressive when they deliver aggressive arguments than non-aggressive arguments, however, there is no evidence that this is gendered. Despite existing stereotypes that women should not be aggressive (Huddy and Terkildsen, 1993a), voters do not perceive arguments by women as any more or less aggressive than equivalently aggressive arguments by men. The middle plot – which shows the emotion findings – shows a very similar result. Voters do perceive emotional arguments to be more emotional than non-emotional arguments, but this again is not gendered. Finally, the bottom-right plot shows that while voters perceive statistical arguments as more evidence-based than anecdotal arguments, there is again no evidence that this effect is gendered. Voters do not perceive anecdotal or statistical arguments delivered by men to be more evidence-based than equivalent arguments delivered by women. Further, I also expected that differential perceptions of styles might serve as a mechanism through which voters differentially evaluate politicians. I find no evidence of differential effects for style perceptions and, consequently, variation in likeability and competence evaluations are very unlikely to be explained by differential perceptions of styles themselves.

In the theory outlined above, I argued that women would be punished when they communicate in ways which are incongruent with traditional communal female stereotypes. However, while I find that the styles politicians use has important consequences for voters’ evaluations, I find no evidence that women disproportionately suffer, at least with respect to likeability and competence evaluations, when they express styles which are incongruent with female stereotypes. Further, I find no evidence

---

I estimate the same model for the competence outcome:

$$\begin{aligned}
 Competence_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 Emotion_j + \beta_3 Aggression_j + \\
 & \beta_4 Statistics_j + \beta_5 Anecdote_j + \beta_6 (WomanMP \cdot Emotion) + \\
 & \beta_7 (WomanMP \cdot Aggression) + \beta_8 (WomanMP \cdot Statistics) + \\
 & \beta_9 (WomanMP \cdot Anecdote) + \gamma X_{i(j)} + \epsilon_i
 \end{aligned}
 \tag{5.6}$$

that voters' perceptions of the styles themselves are gendered.<sup>11</sup>

Finally, although not the main quantity of interest, it is possible that men and women MPs may receive differential evaluations in likeability and competence depending on the policy area in question. As I argue in chapter 4, men and women focus on different topical areas in parliamentary debate and the incentive to do so may in part be driven by internalised pressures from colleagues or voters. In the appendix in section C.7, I analyse whether this is the case. While I find some interesting differences across policy areas – for instance, politicians delivering arguments on the housing and health policy areas are considered to be more likeable than in the transport policy area – these effects do not differ by the gender of the MP. That is, it is not the case that women politicians are particularly likeable when delivering arguments on healthcare.

#### *Conditional effects by voter and MP gender*

How might (gendered) judgements of the styles politicians use vary by voter gender? Prior work examining the importance of voter gender on evaluations of politicians has produced inconclusive results. Some work has uncovered no differences between men and women voters (Bauer, 2015b), while other work finds that gender affinity makes women voters more supportive of women politicians (Sanbonmatsu, 2002), and that voters may be less likely to penalise politicians from their own gender (Rudman and Goodwin, 2004).

To assess heterogeneous effects by voter gender, I carry out two sets of analysis. First, I examine whether men and women voters are equivalently sensitive to politicians' style usage. In figure 5.4, I present the results of analysis where I interact voter gender and style treatment status for each of the outcomes. Only for the evidence style are there differences between men and women voters. While men voters do not find politicians' use of anecdotes as less competent or evidence-based than statistical evidence, women voters do. Further, while the use of anecdotes improves men voters' likeability assessments of politicians relative to the use of statistical evidence, this is not the case for women voters. To the extent that there are differences in how men and women voters evaluate politicians' use of styles, the differences are concentrated among the evidence style. I present the full regression results from this analysis in tables C.14–C.16 in appendix section C.8.

<sup>11</sup>My primary quantities of interest are interaction effects between style treatment status and MP gender that are non-significant. A plausible concern is that the design I present is insufficiently powered to detect the effect sizes reported. To address this concern, in the appendix in section C.9, I report the results of a power analysis where I simulate the data collection process for the fixed sample size available for different hypothetical standardised effect sizes ranging from very small to large effects according to conventional standards. I find that if the true effect size was small then my design may not be sufficiently powered to detect this effect, however the sample size I have is sufficient to detect medium effect sizes.

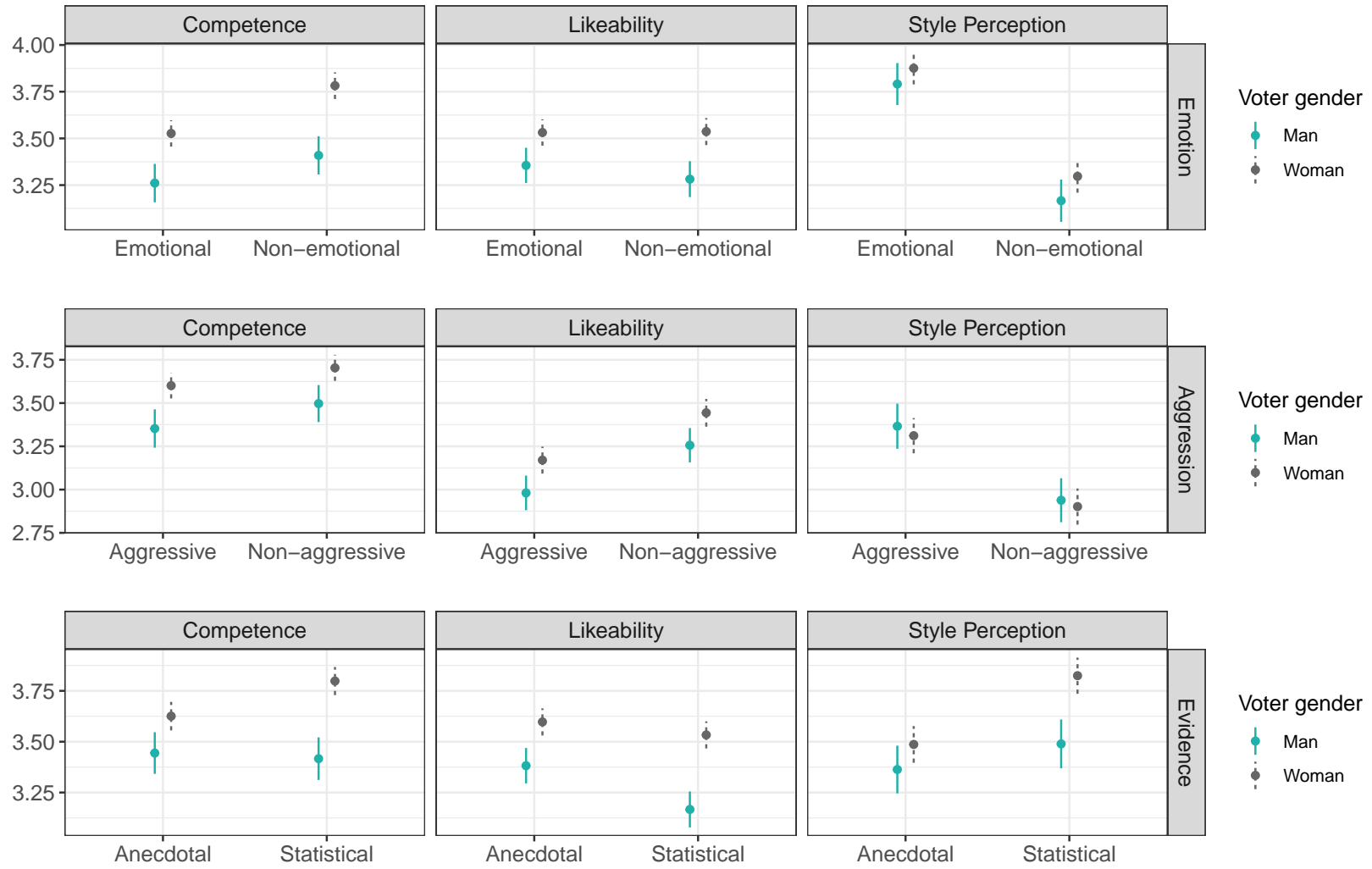


Figure 5.4: Conditional relationship between voter gender, style treatment status, style perceptions, MP likeability, and MP competence evaluations



Second, I assess whether men and women voters are differentially sensitive to the extent to which women politicians conform to stereotype-congruent behaviours. Figure 5.5 presents the results from an analysis where I subset the data into men and women voters and replicate the main analysis described above. The rows represent the styles, and the columns represent the outcomes. Each panel shows the four combinations of voter gender and MP gender, where men voters' evaluations of men politicians are displayed in the dark blue squares, men voters' evaluations of women politicians are displayed in the light blue circles, women voters' evaluations of men politicians are displayed in the triangular yellow points, and women voters' evaluations of women politicians are displayed in the maroon diamonds.

While for aggression and evidence there are no differences, this is not the case for emotion. For likeability, women politicians in particular are rewarded for expressing emotional styles instead of non-emotional styles among women voters; an effect I do not find for men voters. For competence, I find that while women voters find emotional politicians overall to be less competent than non-emotional politicians, they give women MPs *more* of a competency reward than men MPs for expressing emotional styles. I again see no such effect among men voters. Finally, for style perceptions, I see that women politicians are perceived as particularly emotional compared to men politicians when they express emotional styles; an effect I do not see for men respondents. I present the full regression results from this analysis in tables C.17–C.19 in appendix section C.8.

Therefore, women voters give a larger likeability and competence reward to women politicians who are emotional and perceive women politicians as more emotional than men politicians. I find no evidence of these effects among men voters. Consequently, to the extent that there is any evidence of women politicians being rewarded for conforming to stereotype-congruent styles, the effects are concentrated amongst women voters for the emotion style.

## 5.6 Conclusion

Do the styles politicians use influence how voters evaluate them, and does this matter more for women than for men? In this chapter, I address these questions through a novel survey experiment where I present voters in the UK with speeches where the argument style and gender of the politician delivering the argument are varied. This enables me to identify, first, whether politicians experience a backlash effect with respect to evaluations of likeability and competence when they deploy styles that are gender stereotype-incongruent, and second, whether voters' differential perceptions of the styles themselves might explain this backlash.

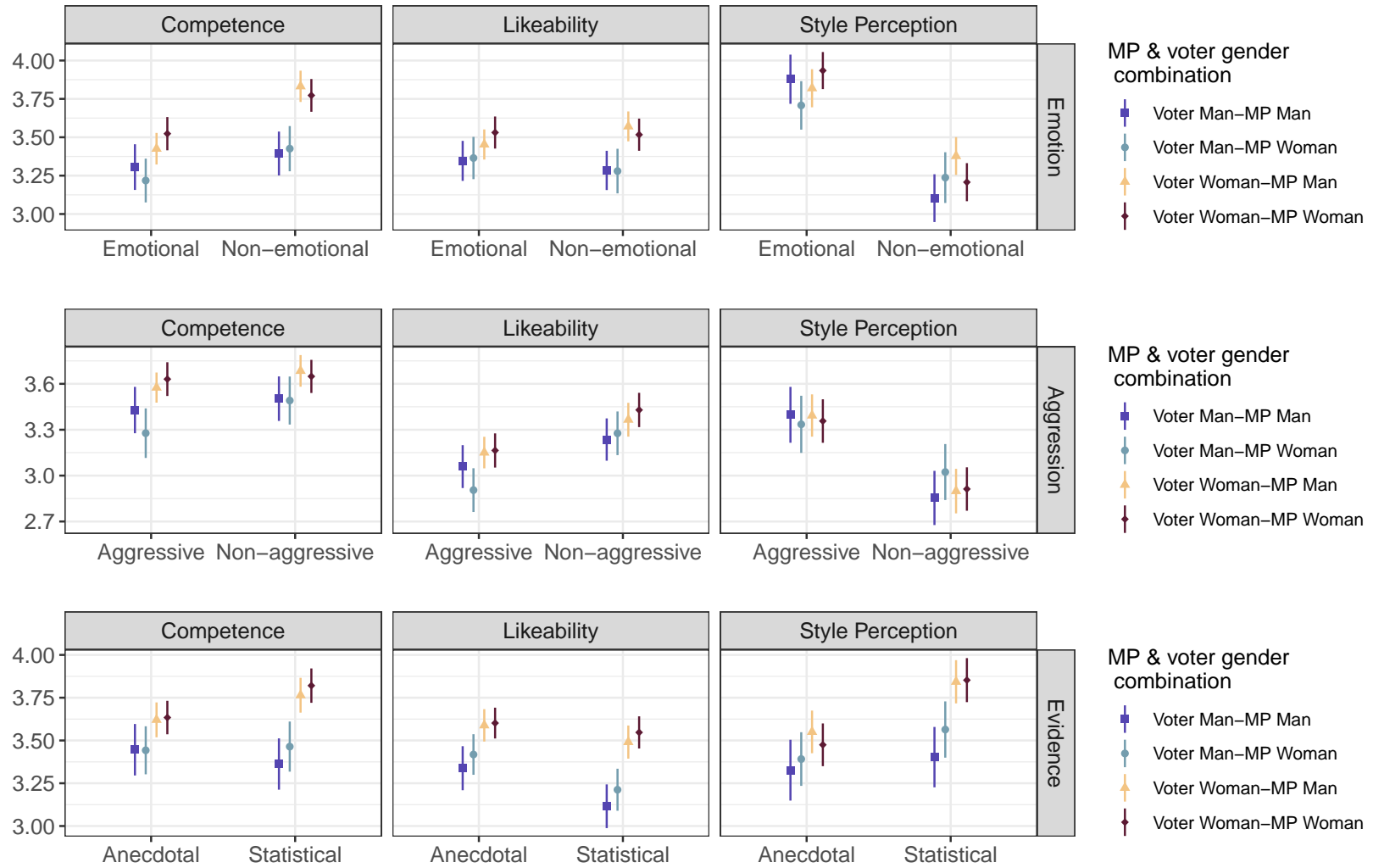


Figure 5.5: Conditional relationship between MP gender, voter gender, style treatment status, style perceptions, MP likeability, and MP competence evaluations

I report four main findings. First, style usage has important consequences for voters' evaluations of politicians. Politicians who are unaggressive and draw on anecdotes are more likeable, whereas politicians who are unemotional and refer to statistical evidence are more competent. Second, I find no evidence that voter evaluations of politicians are gendered. In particular, women politicians are *not* punished for stereotype-incongruent behaviour. Third, while there is clear evidence that voters can identify the styles politicians use, I also find no evidence that voters' perceptions of the styles themselves are gendered. Gender bias in voters' perceptions of the argumentation styles themselves is very unlikely to explain variation in the likeability and competence evaluations. Fourth, I find some evidence that this differs by voter gender.

Across the various styles and outcomes, the main finding I document is therefore that styles influence voters' evaluations of politicians, but these evaluations do *not* vary by MP gender. Why do I find little evidence of gender bias in voters' evaluations of women politicians, and what implications may these findings have for voting behaviour? First, one possible explanation for the "null" effect for the style perceptions might be that there are competing expectations about whether stereotype application means voters perceive politicians to be *more* or *less* like the stereotypes relevant to their gender (Koch, 2000). To illustrate this tension, let's take the example of aggression (Brescoll and Uhlmann, 2008). If a voter is presented with an aggressive argument, one outcome is that they will perceive aggression as more prevalent when delivered by a *man* as this confirms stereotypes about men and aggression (Koch, 2000). However, an alternative outcome is that a voter might rate aggression as more prevalent when delivered by a *woman* as this violates stereotypes about women and aggression (Cassese and Holman, 2018). The former implies a confirmation bias: voters pay attention to information that confirms a stereotype. The latter implies a backlash effect: voters pay attention to information that violates a stereotype. While the findings I present do not support either perspective in isolation, it is possible that there is some degree of backlash *and* confirmation bias at play which results in the "null" effect because of these offsetting heterogeneous dynamics.

Second, while I find that style usage influences voters' evaluations of the likeability and competence of politicians, these findings do not, however, enable me to assess whether these evaluations have downstream consequences for voting behaviour. The styles and personalities of leaders have long been considered an important determinant of voters' attitudes (Declercq, Hurley and Luttbeg, 1972). Further, as partisanship in the electorate has declined over time (Dalton and Wattenberg, 2000), and voters have become increasingly volatile (Cutts et al., 2020; Fieldhouse et al., 2019), then the styles politicians express, and associated evaluations, may have even increased in importance as determi-

nants of vote choice as voters base their decisions on factors beyond party. Further, while I cannot directly assess whether these evaluations influence vote decisions, previous UK-based work has shown that voters' evaluations of politicians' competency do influence their voting preferences (Green and Jennings, 2017). It is therefore not unreasonable to assume that likeability and competence evaluations may inform vote intention.

My findings suggest that politicians face trade-offs in evaluations: while styles compatible with communality lead to positive likeability evaluations, styles compatible with agency lead to positive competency evaluations. Should politicians prioritise competency or likeability evaluations? Traditional accounts of leadership have suggested that competence is important in informing vote choice (Green and Jennings, 2017), and congruent with the traits and behaviours deemed necessary and suitable for leaders (Elgie, 2015; Koenig et al., 2011). According to traditional accounts, we may consider competence the more important evaluation to optimise. However, a trend that is common across many political contexts in recent decades is that voters are increasingly dissatisfied with politics and find that politicians are out-of-touch and unlike "normal people" (Clarke et al., 2018; Garzia, 2011). In response to voter dissatisfaction with this type of politics and politician, there is an increasing desire for politicians who are instead human, personable, charismatic, engaging, and in-touch (Meyrowitz, 1985; Valgarðsson et al., 2021). The move to telegenic, human, and personable styles has been said to be a core element of the success strategies of populist candidates (De Vries and Holbolt, 2020), and recent examples of the use of these styles, such as the widespread fame and popularity of Ukrainian President Volodymyr Zelensky, suggests that voters and the media find this to be compelling.<sup>12</sup> While traditional accounts suggest that ensuring competency is more important in determining candidate success, voters have begun to place increasing importance on the likeability of politicians.

As far as I am aware, we currently lack systematic evidence on which evaluations of politicians among the many that prior work has studied, such as likeability, competence, honesty, hard-workingness, or charisma, matter most in informing voting behaviour. A fruitful avenue for future work would, therefore, be to identify which traits voters prioritise in their decision at the ballot box. Such a study would be useful not only for understanding the wider implications of the results I present here, but also for the plethora of experimental studies that assess how a variety of features of politicians' behaviour – such as whether they are corrupt (Eggers, Vivyan and Wagner, 2017), loyal to their party (Campbell et al., 2019a), or their representational efforts (Blumenau, Wolkenstein and Wratil, 2022) –

---

<sup>12</sup>See, for example, "The Zelensky Effect: How To Engage, Energize and Unleash Your Organization's Potential", *Forbes*, 29th March 2022.

influences how voters evaluate them.

Finally, at the core of the idea that women politicians face double standards when they violate stereotypically expected behaviours is that voters actually hold these expectations for women's behaviour in the first place. However, studies show that the public's perception of the validity of these stereotypes has shifted over time. Women have been seen as increasingly agentic over time (Eagly et al., 2020; Sendén et al., 2019). Voters in the UK have also become more gender-egalitarian in their attitudes over time (Taylor and Scott, 2018), and have become markedly less likely to support traditional gendered divisions in social roles (Shorrocks, 2018). Further, the evidence from chapter 3 suggests that UK politicians have also come to behave in a way that is less consistent with traditional gender stereotypes. Women politicians have decreasingly used communal styles – including anecdotes and emotion – and increasingly use agentic styles – including aggression and statistics – over time. The pessimistic assumption is that this behaviour change might be met by backlash from voters. However, the results presented here suggest that this may not materialise, as UK voters do not seem to unjustly penalise women politicians for stereotype-incongruent behaviour.

Of course, without a study from 20 years ago to compare these findings to, it is not possible to know whether UK voters in previous eras did apply these descriptive stereotypes or punish women politicians for stereotype-incongruent behaviour. Yet, if voters no longer hold the same stereotypical expectations about men's and women's behaviour, and politicians decreasingly behave in accordance with traditional stereotypes, it may not be surprising to uncover that women politicians are not punished for behaviour that violates traditional stereotypes.

## Chapter 6

# Conclusion

Do politicians behave in accordance with traditional gender stereotypes? Does the pressure to do so shift by context and over time? Do voters uphold these stereotypes when they perceive and evaluate politicians' behaviour? Focusing on voters and elites in the UK, this thesis has provided answers to these three questions. In answer to the first two research questions, I find that politicians do behave in accordance with gender stereotypes, but that this varies both over time and over the course of politicians' careers. In answer to the third research question, my results suggest that voters apply stereotypes when evaluating politicians' behaviour – for instance, voters see politicians using “communal” styles, such as emotionality or drawing on anecdotes, as more likeable and those using “agentic” styles, such as unemotionality and drawing on statistical evidence, as more competent – but that men and women politicians are not held to different standards for these behaviours. Taken together, this thesis provides important theoretical arguments and empirical evidence concerning the *dynamic* validity of stereotypes for describing elite and voter behaviour in the UK.

In this final chapter, I summarise the findings of each paper, consider how these findings relate to one another, and discuss what the wider implications of this research are for politicians, voters, and political institutions. I then turn to a discussion of the generalisability of these findings, consider some of the limitations, and conclude by highlighting some potential avenues for further research.

### 6.1 Key findings and substantive contributions

Existing accounts posit that politicians face clear incentives to behave in ways which are consistent with gender stereotypical norms. These incentives are said to induce differences in the ways in which men and women carry out their political roles with respect to the issues they prioritise when participating in political debate and the style in which they deliver these arguments (Bear and Woolley, 2011; Eagly, 2013;

Herrnson, Lay and Stokes, 2003; Huddy and Terkildsen, 1993a; Karpowitz and Mendelberg, 2014). Central to the incentive to behave in a way that is consistent with stereotypes are expectations about how voters reward or punish politicians for their behaviour during campaigns or once in office (Brooks, 2013; McDermott, 1997; Schneider and Bos, 2019). The conclusions from previous work suggest that gender stereotypes inform both the ways that politicians behave in office and inform voters' judgements about this behaviour. I have set about assessing whether stereotypes are influential in shaping the behaviour and attitudes of politicians and voters in the UK.

Chapter 3 investigated whether stereotypes serve as accurate descriptions of the styles men and women express in parliament. Building on work that considers stereotypes to be dynamic concepts (Diekman and Eagly, 2000; Eagly and Wood, 2012; Eagly et al., 2020), I argued that between 1997 and 2019, my period of study, UK MPs conformed less to gender stereotypes now than in the past, and that women in particular have increasingly adopted styles that are consistent with masculine "agentic" stereotypes of communication over time. To evaluate this argument, I focused on politicians' styles as expressed in parliamentary debates and developed novel quantitative text analysis techniques to measure a diverse set of styles at scale. These were: human narrative, affect, positive emotion, negative emotion, factual language, aggression, complex language, and repetition. My goal was to construct measures that closely approximated the conceptual definitions of each of the styles I identified in my review of the literature. I highlighted a common limitation of existing approaches to detecting style that suffer from a well-known problem for off-the-shelf dictionaries: the words that demonstrate a given concept in one context may be poorly suited to detecting the use of that concept in another context. I circumvented this problem by using a word-embedding model which enabled me to capture how each of the styles manifest in the specific context of UK parliamentary debate.

Applying these measures to nearly half a million speeches delivered in the House of Commons between 1997 and 2019, I reported three main findings. First, at the aggregate level, I documented gendered patterns of style that are broadly consistent with the conclusions from previous literature. Women politicians' speeches are marked by higher levels of emotionality, positivity, and make greater use of human narrative, while men make greater use of aggressive and complex language. At the aggregate level, my findings suggest that politicians do behave in a way that is broadly consistent with gender stereotypical expectations. Second, and crucially, I find that these differences have reduced dramatically over time. While in the late 1990s and early 2000s men and women did express communicative styles that are broadly consistent with gender stereotypes, I show that these differences have reduced dramatically over time. Across almost all the styles I studied, there has been a convergence

in men and women's usage over time. For cases where I report diverging behaviour over time, these stylistic shifts run counter to gender-role expectations. Third, I show that the evolving variation in style use has primarily resulted from women's decreasing use of communal styles, and increasing use of agentic styles, over time.

While most of the literature on gender differences in legislator behaviour has considered the accuracy of stereotypes at a single point in time, I provide an important contribution to this work by considering the influence of stereotypes over time. In doing so, I show that stereotypes are now less accurate at describing politicians' behaviour than they were in the past. While it may be true in some settings that politicians behave in a way that is broadly consistent with stereotypes, my research shows that gender stereotypes of communication styles have become significantly less predictive of the reality of contemporary legislator behaviour.

Chapter 4 assessed whether stereotypes guide the sets of issues that politicians raise and prioritise during their time in office, and, in particular, how this could change over the course of politicians' careers. I argued that gender differences in the extent to which women politicians raise issues traditionally associated with women would be most pronounced among junior politicians, however this difference would decrease with increased parliamentary experience. To test this argument, I leveraged quantitative text analysis approaches to develop new measures to capture the sets of issues that politicians raise. Following recommendations from previous literature about the appropriateness of determining *a priori* which issues should be associated with men and women (Celis et al., 2014; Yildirim, 2021), I focused on a range of issues that has previously been identified as stereotypically "feminine" and "masculine" (Huddy and Terkildsen, 1993a), but also a wider set of political issues that are of central importance to UK politics (Bevan and Jennings, 2019) to assess gendered dynamics on a wide range of policy areas. Further, I assessed the degree to which politicians amplify women voters' voices in legislative debates.

I reported four main findings. First, at the aggregate level, the topics that men and women participate in are broadly consistent with the traditional divisions highlighted in prior literature. Women politicians dedicate more attention to raising issues that women are traditionally expected to have greater expertise in and authority on, such as children and family, health, education, and welfare. While for some issue area on which men are expected to focus more attention, such as the economy or trade, I uncovered no gender differences, men do raise matters of defence policy more than women. This evidence is, therefore, broadly consistent with conclusions from previous work. Second, these aggregate gender differences masked significant variation over the course of politicians' careers. Among



junior politicians, women talk significantly more about traditionally “feminine” policy areas, however this gender gap reduces markedly among senior politicians. Further, this pattern was concentrated only among these policy areas, and not among the wider set of issues politicians raised. Third, I found that women raise women’s experiences to a greater extent than men, and that this difference does not change across the career cycle. Finally, I found that this difference held across debates on all policy areas and that this, too, does not shift with political experience. Women therefore diversify the sets of policy areas they focused on the longer they are in parliament, however they also continue to raise women’s perspectives and experiences in this wider set of political issues.

While existing investigations into the issues that politicians raise in the UK context have examined these questions in a single parliamentary session (see [Bird, 2005](#), on the 1997-98 session; and [Catalano, 2009](#), on the 2005-07 session), my paper represents the first investigation over a long period of time into the sets of political issues that men and women politicians in the UK House of Commons champion. Further, I consider how the sets of issues that politicians dedicate their attention to may shift over the course of their careers. I contribute to a growing body of literature that has sought to understand how incentives shape politicians’ representational efforts ([Bailer et al., 2022](#); [Höhmnn and Nugent, 2021](#); [Weeks et al., 2022](#)), by arguing that incentives shift over the course of politicians’ careers because of the experience and credibility they gain throughout. As such, my results have positive implications for both the experiences of women politicians in office, but also for the substantive representation of women’s interests in politics more broadly. As women politicians become more senior, they can champion a wider variety of issues beyond those traditionally associated with women, while continuing to raise women’s experiences and perspectives in these issues, thus ensuring that the interests of women voters are raised in debates across the policy-making process. These findings contribute to a wider body of literature that suggests not only is it important to incorporate more women numerically into the policy making process ([Wängnerud, 2009](#)), but also that women remain in politics for longer ([Childs and Krook, 2009](#); [Humphreys, Masters and Sandbu, 2006](#); [Wahman, Frantzeskakis and Yildirim, 2021](#)). Work by [Blumenau \(2021c\)](#) demonstrates a role model effect, whereby the appointment of women to leadership positions in the UK House of Commons increases both the participation of other women politicians and their influence in parliamentary debates. The implication of these results is that senior women can boost the presence and power of women within the legislature. My results complement these findings and suggest that ensuring there is effective substantive representation of women’s experiences, priorities, and perspectives in the policy process can be achieved by not only increasing women’s presence in politics, but also ensuring that women reach sufficiently senior levels within it.

A key implication of the findings from chapter 3 is that gender stereotypes with respect to *how* politicians talk have become less accurate over time. Might gender differences in the traditional divisions of issues politicians raise shifted over time as well? I argued that changes to the social roles played by women in public life, and in politics, may have reduced the content and validity of stereotypes in the eyes of the public. We might imagine that voters now have less clear expectations in the sets of issues men and women are perceived to have expertise in, and may now be less likely to sanction women politicians for stereotype-inconsistent behaviour. However, as I show in appendix B.8, while there are some year-by-year changes in the gender gaps for some topical areas, the differences in the attention men and women politicians bring to different topical areas follow no clear trends over time. Conditional on experience, women continue to talk consistently more about the “feminine” policy areas than men between 1997 and 2019. My results therefore suggest that while the style of politicians’ speech has diverged from stereotypical trends over time, the sets of issues they raise have not. While women may not focus less on stereotypically “feminine” policy areas over time, they do focus less on these issue areas over the course of their careers. These results suggest that it is changes that occur over the course of politicians’ careers, such as their increased experience in a diversity of policy areas and credibility in the eyes of their colleagues, that influence the issues they raise as opposed to factors outside of parliament. While women remain underrepresented with respect to numbers and tenure in parliament, the numbers of women have been increasingly steadily over time in the UK. I show that as women become increasingly senior, they focus less on stereotypically feminine policy areas. As women continue to enter parliament in growing numbers over the coming years, and eventually become increasingly senior, then, in the future, we may witness a divergence in the extent to which politicians raise stereotypically gendered policy areas. While I can only speculate about the future of gendered behaviour change, my results emphasise the importance of continuing to address these questions in future scholarly efforts.

Chapters 3 and 4 show the validity of *descriptive* stereotypes for informing politicians’ behaviour is contingent on time and seniority. Chapter 5 considers instead the extent to which voters uphold *prescriptive* stereotypes when engaging with politicians’ behaviour. Specifically, I focused on whether voters are biased in their perceptions and evaluations of the ways in which politicians communicate, and, consequently whether evaluations of elite political communication are gendered. To do so, I focused on styles for which the gender stereotypes literature outlined clear expectations about men and women’s behaviour, and four out of the eight styles that I focused on in chapter 3. Work to date examining how voters evaluate stereotype-(in)congruent behaviour has tended to focus on isolated

behavioural traits (e.g., Bauer, 2015a; Cassese and Holman, 2018; Krupnikov and Bauer, 2014). Instead, I made progress here by focusing on how voters evaluate politicians' use of a diversity of styles consistent with both feminine "communal" stereotypes and masculine "agentic" stereotypes. By doing so, I was able to assess whether there are asymmetric standards in the degree to which men *and* women politicians are punished for stereotype-incongruent behaviours. I argued that first, voters would perceive the styles that men and women express to be different based on stereotypical expectations of style usage, second, women politicians would incur backlash from voters when violating gender-based norms, and third, that differential perceptions of the styles themselves would serve as a potential mechanism through which voters' gendered evaluations of politicians might become manifest.

I reported four main findings. First, I found that politicians' style usage has important consequences for how voters evaluated them. Politicians are deemed to be more likeable when they are emotional and draw on anecdotes, however, they are regarded as more competent when they are unemotional and make use of statistical evidence. Second, contrary to my own expectations, and the expectations from the gender stereotypes literature, I found no evidence that these evaluations were gendered. That is, while all politicians are evaluated as less likeable when they are aggressive, there is no evidence that women in particular incur negative evaluations. Third, while I showed that voters were able to accurately identify and characterise the styles that politicians used, voters' perceptions of the styles themselves are not gendered. Voters do not perceive unemotional arguments by women politicians to be any more or less emotional than unemotional arguments by men, nor do they perceive aggressive arguments by women to be any more or less aggressive than equally aggressive arguments by men. Finally, I found some evidence that these evaluations differed by a voter's gender. Women voters reward women politicians for stereotype-congruent behaviour: they gave a larger likeability and competence reward to women who are unemotional compared to when they are emotional. Further, when assessing the emotionality of an argument, they perceive arguments delivered by women politicians to be more emotional than equivalent arguments delivered by men. These effects are concentrated only among women voters and not men.

While there is a rich literature on the influence of gender stereotypes on voter *evaluations* of politicians, this thesis is the first to address whether voters differentially *perceive* politicians' argumentation styles based on gender alone. This enabled me to investigate one potential mechanism through which differential evaluations might become manifest: if voters perceived women compared to men as, say, particularly aggressive even in the absence of any objective differences, this may explain why women in particular have been shown to suffer with respect to likeability when they are aggressive. Instead, I

found no evidence that voters' perceptions of the styles themselves are gendered, and, consequently, variation in likeability and competence evaluations are unlikely to be explained by differential perceptions of the styles themselves. Further, a commonly held perception is that men's leadership styles are preferred for office than women's (Fox and Oxley, 2003; Jalalzai, 2016), and that when women try and adopt these styles, they violate feminine stereotypes, and will lose out. However, the findings I present suggest that the reality for women politicians is more positive than this, as the expectation that they must avoid these styles for fear of negative backlash from voters is somewhat misguided. While, of course, women politicians may be sanctioned for stereotype-incongruent behaviour from other sources, the findings I present here suggest that voters do not always punish women in the way that common theories of gender stereotyping may have predicted, at least with respect to written speeches.

Taking these findings together, the conclusions from chapters 3 and 4 emphasise the importance of adopting a *dynamic* understanding of the influence that gender stereotypes wield over men and women's behaviour. Past literature has tended to consider these questions in rather a static manner and has presented results which suggest that stereotypes are accurate descriptions of politician's behaviour. Indeed, my own aggregate-level findings are consistent with these conclusions too. However, when considering how conformity to stereotypes shifts both over time and throughout the course of politicians' careers, I show that the validity of these stereotypes is far from fixed, and that the degree to which they accurately describe politicians' behaviour varies markedly over time and across contexts. Approaches that report only aggregate-level assessments at a fixed point may serve to confirm stereotypical expectations and further encourage the continued upholding of such stereotypes by the public. That is, by assessing whether there are gender differences at only a particular point in time, and uncovering stereotype-consistent patterns, we risk perpetuating the idea that stereotypes are accurate and do not assess how stereotypes are likely dynamic and subject to change as societal gender roles and attitudes continue to shift. Instead, my findings challenge the validity of these descriptive stereotypes. Prescriptive stereotypes for how men and women *ought* to behave are ultimately rooted in our collective understanding of how we expect men and women *will* behave (Eagly and Wood, 2012; Gill, 2004). It is important to document when behavioural shifts run counter to gender-based stereotypes, as these conclusions may help to further update voters' stereotypical priors and, in turn, further diminish the degree to which politicians are subject to penalties for failing to conform to stereotypical expectations.

The findings from chapter 5 also challenge the widely held expectation that voters in the UK uphold gendered behavioural prescriptions when evaluating politicians. An implication of the findings

from all three papers is that, in the contemporary UK context, stereotypes are now less accurate for describing behaviour in the legislature and that voters in 2021, at the time the experiment was fielded, do not judge politicians based on these stereotypes. The consistency in these findings therefore suggest that, at least with respect to gender stereotypes as I have studied them, stereotypes are less influential now than they once were. While voters may not punish women for expressing styles that are stereotype-incongruent, my results suggest that the styles politicians use have important implications for how voters evaluate them. In particular, I show that styles associated with feminine “communal” stereotypes lead voters to evaluate politicians as *more* likeable but *less* competent. While voters may no longer punish women for stereotype-incongruent styles, my results suggest that voters do penalise politicians with respect to competency evaluations and reward politicians with respect to likeability evaluations when they express styles which are traditionally associated with women’s behaviour. Women politicians may not be disproportionately penalised, however, the stylistic features associated with the traditional “feminine” gender roles may still be punished. Voters may therefore still be aware of the descriptive stereotypical categories we associate with masculine and feminine behaviour and hold an outright preference for traditionally masculine behaviours with respect to assessing politician competence.

Given men’s traditional occupation in leadership positions and political office more broadly, and, as a result, the congruence between masculine stereotypes and leadership stereotypes, it may not be that surprising to uncover that politicians are deemed to be more competent – a characteristic that has historically been deemed suitable and necessary for politicians (Elgie, 2015; Foley, 2013; Koenig et al., 2011) – when expressing masculine agentic behaviours. Recent work has shown that those who run for office, compared to those who do not run, express behaviours that are more consistent with agentic stereotypes, and less consistent with communal stereotypes (Conroy and Green, 2020). Overall, therefore, the traditionally masculine ethos of politics may still favour agentic behavioural traits with respect to ensuring politicians are seen to be competent. However, my findings, which show both that women politicians are increasingly behaving in a way that is consistent with agentic stereotypes over time *and* that voters do not punish women politicians any more than men for expressing agentic styles, suggests that, today, women lose out less for this behaviour than they did in the past.

Together, the empirical results from each of the three papers which form the core of this thesis provide several important contributions to the study of the influence of gender stereotypes in politics. While distinct in question and approach, a common theme to each of the findings is that the influence that gender stereotypes can wield over politicians and voters is complex, can change according to

time and context, and is subtler than traditional accounts would suggest. I turn next to considering the wider implications that the findings I present here have for politics and policy.

## 6.2 Implications for politics

The research undertaken in this thesis has important implications beyond academic accounts of gender, stereotypes, political behaviour, and parliament. First, the COVID-19 pandemic saw rise to the widely popularised claim that women leaders have a style that is systematically different, and even “better suited” to handling the pandemic than men. Media outlets were largely responsible for the popularity and far-reaching dispersion of this claim, and published reports that women-led countries responded faster and more effectively communicated their pandemic policies emerging in *The Guardian*, *Vogue*, *The New York Times*, *The Washington Post*, *CNN* and many others. In particular, women leaders were said to operate with leadership styles marked by empathy, compassion, care, clear communication, and to have “unique talents and capabilities” that enabled them to handle the crisis effectively. While academic enquiries have largely dismissed the suggestion that leader gender was responsible for pandemic performance (Piscopo, 2020), primarily because of a host of confounding factors relating to the types of countries that elect women to high office in the first place and their capacity to respond to crises, the claim that women have a unique and different leadership style that can fix political issues persists.

The idea that women politicians offer a behavioural style that is different and perhaps better suited to tackle some of the less attractive features of our political systems has been often cited in the UK case specifically.<sup>1</sup> Indeed, there is a strand of popular commentary that suggests that the election of more women into Westminster would ameliorate some of the more adversarial, aggressive, and unattractive features of parliamentary culture. However, my results cast doubt on the idea that increasing the number of women would result in a form of politics that is kinder, gentler, and more pleasant, as this assumption is based on an outdated assumption about the distinctiveness of women MPs’ political styles. Over the period I studied, I showed that while men’s use of human narrative, emotion, and positive emotion has remained relatively fixed, women have made decreasing use of these styles over time. Further, while men’s use of negative emotion has stayed reasonably stable and their use of aggression has declined slightly, women’s use of negative emotion has increased notably and their use of aggressive language increased slightly over time. The results I present suggest that

<sup>1</sup>See, for example, “Designing a new parliament with women in mind”, *Democratic Audit*, 29th July 2016; “More women MPs may be the only way to reduce sexism in parliament”, *Sky News*, 1st May 2022; “Woman’s Hour”, *BBC Radio 4*, 7th July 2022.

the election of more women alone may not be sufficient to significantly “better” parliamentary culture.

Further, by placing hopes of change of this scale on the shoulders of women, proponents of these views may be setting women up to fail if their increased presence does not improve political culture. It would be better to pursue the much-needed effort of increasing women’s presence in politics and leadership on the basis of equality, fairness, representation, responsiveness, and ensuring that a diverse range of experiences, perspective, interests, and priorities are given space in the policy-making process. Indeed, the results I present in chapter 4 suggest that women politicians do more to raise the perspectives of women in legislative debates, that they do so in debates on a wide range of policy areas, and that they continue to do so throughout their parliamentary careers. Increasing the number of women in politics should undoubtedly be an important normative goal of our political institutions. However, if we want to strive for more effective approaches to tackling political crises, and reach a more deliberative style of politics, then we need to look towards meaningful reforms to do so,<sup>2</sup> instead of simply hoping that the election of more women will miraculously better our political culture.

Second, an oft-cited criticism by the public of politicians’ behaviour in the UK House of Commons, and in particular in Prime Minister’s Questions, is that it is akin to “playground arguing” (Hansard Society, 2014, 20), that it is “gladiatorial” (Lovenduski, 2012, 319), and at times, similar to a “Punch-and-Judy style of politics”. Prior academic accounts suggest that this sort of aggressive and adversarial behaviour is unpopular with the public (Dahlerup, 2017), as voters dislike uncivil candidates (Bauer, Kalmoe and Russell, 2022). Further, compared to other approaches to political argumentation techniques, rhetoric that invokes attacks towards others is regarded to be relatively unpersuasive (Blumenau and Lauderdale, 2022b). My results suggest that politicians who behave in a way that is angry and aggressive are viewed by voters to be neither competent nor likeable, that this holds when judging the behaviour of men and women politicians alike, and that these views are shared by men and women voters. Coupled with this, while I measure the styles of politicians up until 2019, at the end of my sample period, I show that while men have historically made greater use of aggressive language, women politicians may even be *increasing* their use of aggression over time to match the levels of men. This suggests a trend where, if anything, the use of aggressive language – which is viewed poorly by voters – could even be increasing over time.

Voters are increasingly shown to be disengaged and disinterested from politics and politicians (Clarke et al., 2018; Russell, 2005), and have become more volatile in their voting behaviour (Fieldhouse

---

<sup>2</sup>See, for example, “Equality in the heart of democracy: A gender sensitive House of Commons”, *Women and Equalities Committee*, 23rd February 2022.

et al., 2019). Further, recent survey-based evidence suggests that voters in the UK have very little trust in political elites, and prefer politicians to instead act with integrity, honesty, and to be in touch with ordinary people (Renwick et al., 2022). Therefore, as voters increasingly express dissatisfaction with politics and political institutions, my findings contribute to a wider body of evidence that suggests that politicians may be behaving in a way that potentially puts voters off politics. By presenting new evidence that voters, men and women, do not find aggressive behaviour from political elites to be engaging nor congruent with competence, my results emphasise the importance of politicians adopting different behaviours if they want to seriously address voter dissatisfaction with politics and politicians.

Third, while my findings suggest that the influence stereotypes wield over politicians' and voters' behaviour and attitudes may be more complex and dynamic than previous accounts suggest, these findings do not imply that stereotypes play no role in politics. The research in this thesis makes an important contribution by providing new insights on three crucial ways in which stereotypes may be influential: on the manner in which politicians debate in parliament, on the sets of issues they raise, and how voters engage with this behaviour. My results do not, however, enable me to assess the extent to which the media stereotypes women politicians. Existing accounts suggest that men receive more coverage than women (Banwart, Bystrom and Robertson, 2003; Bystrom, Robertson and Banwart, 2001; Falk, 2008), yet, when women are covered, this coverage tends to be both more negative (Gidengil and Everitt, 2003; Heldman, Carroll and Olson, 2005; Kahn, 1992; Ward, 2016), and to understate women's abilities relative to men's (Bauer and Taylor, 2022). Certainly, women politicians in the UK report that sexist media coverage is something they experience.<sup>3</sup> Therefore, voters do not directly stereotype women, but the media may cover women in a way that is both more negative and perpetuates gendered stereotypes. In turn, as the media is the primary method through which voters receive information about politicians, biased media coverage may still inform voters perceptions and evaluations.

Further, my research does not enable me to make statements about the stereotyping that women can incur from those *within* the legislature. Research into gender stereotyping in workplaces more broadly presents considerable evidence that men and women are held to different standards when their competence and leadership abilities are considered (Brescoll, 2016; Eagly and Carli, 2007; Heilman, 2001). As an institution dominated by men, and one that perpetuates masculine norms of behaviour, women can suffer under the pressure of different workplace expectations (Anzia and Berry, 2011; Childs,

<sup>3</sup>See, for example, "What the Angela Rayner story tells us about misogyny in Westminster", *The Guardian*, 26th April 2022; "Women in politics face 'daily' abuse on social media", *BBC News*, 24th May 2020; "Sexist coverage of Liz Kendall and female politicians is insidious and demeaning", *The Guardian*, 20th July 2015; "Jo Swinson says she faces 'sexist' coverage but urges female politicians to get 'stuck in'", *The i*, 3rd December 2019; "Diane Abbott: misogyny and abuse are putting women off politics", *The Guardian*, 14th February 2017; "How the media undermine women political leaders", *Policy Options Politiques*, 12th September 2018.



2004a), and experience more scrutiny for their behaviour than their men colleagues do. Indeed, recent work has shown that women politicians express lower levels of satisfaction with their workplaces (Allen, Cutts and Winn, 2016), that women report experiencing higher levels of emotional labour than men (Weinberg, 2021a), and experience greater pressures to perform, display higher levels of anxiety, report incurring more negative treatment, and receive lower levels of positive feedback than men (Erikson and Josefsson, 2019). Beyond academic accounts too, there are countless examples of women politicians reporting that their behaviour receives more attention and scrutiny than their men counterparts.<sup>4</sup> In short, women MPs still experience stereotyping from their colleagues in parliament and enduring a wider culture of sexism and misogyny.

My research contributes to part of a wider public and academic debate about the ongoing influence of stereotypes and gender bias in the political realm. Sexism undoubtedly remains a problem: politicians report encountering sexist experiences, women remain underrepresented in legislatures and leadership positions globally, and the fleet of recent stories emerging from Westminster specifically on the conduct and behaviour of numerous men politicians showcases that this is the case.<sup>5</sup> My work highlights some areas in politics where stereotyping and gender bias is now less powerful and influential than they once were. These findings should motivate researchers to turn their attention to identifying other forums through which pernicious stereotyping and gendered bias persist.

### 6.3 Generalisability, limitations, and avenues for future work

In this final section, I conclude by considering how these findings may generalise beyond the UK context, some of the key limitations, and propose how these may be addressed in future research.

My focus is on elite and voter behaviour in the context of the UK specifically. The UK political environment represents a valuable setting in which to address my key questions of interest and contributes to a mostly US-based body of literature which has studied the power and influence of stereotypes. Nevertheless, a single country analysis of politician and voter behaviour may not be representative of these dynamics in countries beyond the UK. At the core of the idea that politicians are incentivised

<sup>4</sup>See, for example, “Women MPs say abuse is forcing them from politics”, *BBC News*, 31st October 2019; “Women in Parliament will never feel valued until sexist culture changes”, *UK in a Changing Europe*, 18th May 2022; “From ‘Legs-it’ to ‘calm down, dear’: six times UK MPs have faced sexism”, *The Guardian*, 25th April 2022; “Parliament and misogyny: female MPs still face toxic culture”, *Stylist*, 29th April 2022; “What the hell is wrong with Westminster?”, *Politico*, 29th April 2022.

<sup>5</sup>See, for example, “‘It’s rife’: female MPs tell of climate of misogyny in Westminster”, *The Guardian*, 29th April 2022; “The sexist culture in Westminster reveals itself once more”, *The New Statesman*, 28th April 2022; “Is Westminster a safe place for women to work? Not in my experience”, *Open Democracy*, 19th May 2022; “Britain’s Parliament is Rocked by Sexist Episodes. Again.”, *The New York Times*, 3rd May 2022; “UK Conservative MP ‘watched porn’ in House of Commons”, *Politico*, 27th April 2022; “Sex, booze, and politics: inside the corridors of power in the week that reignited Westminster’s #MeToo”, *The Times*, 30th April 2022.

to behave in stereotype-congruent ways and that women politicians face double standards when they violate stereotypically expected behaviours is that voters hold behavioural expectations in the first instance. In chapter 3, I argued that politicians will have weaker incentives to conform to stereotypes now than they did in the past. To do so, I drew on literature that shows that women in the public have adopted increasingly agentic behaviours over time (Twenge, 2001; Leaper and Ayres, 2007), that the public's perceptions of the validity of gender stereotypes has shifted (Eagly et al., 2020), that voters are less likely to support traditional divisions in gender roles (Shorrocks, 2018), and that there have been dramatic shifts to the roles that women play in political life over time (Blumenau, 2021c; IPU, 2022). While these behavioural and attitudinal shifts have occurred in the UK, the studies upon which I draw suggest that these trends are also in operation in a wide set of context and countries beyond the UK too. The mechanisms upon which my argument relies seem highly relevant to contexts beyond the UK. Inevitably, countries beyond the UK may be at different points on the trajectory of strengthening the power of women in politics and society, as such I encourage future scholars to assess whether politicians in other legislatures shift their adherence to stereotype-congruent behaviours over time.

The conclusions I present in chapter 5 suggest that voters in the UK do not punish women for stereotype-incongruent behaviour. While my findings run counter to the results from US-based work that shows voters often do stereotype politics (e.g., Anzia and Bernhard, 2022; Bauer, Harbridge Yong and Krupnikov, 2017; Cassese and Holman, 2018; Holman, Merolla and Zechmeister, 2016; Krupnikov and Bauer, 2014; Mo, 2015), they are consistent with the small body of work that has addressed the influence that stereotypes may have in the UK context. Mine is not the first UK-based experimental study that has found that the same biases observed in the US do not necessarily travel to the UK (see Eggers, Vivyan and Wagner, 2017; Hargrave and Smith, 2022; Saha and Weeks, 2022; Smith, 2019). One possible explanation might be that the UK has a different experience with gender politics: gender equality is a less polarising and party-political issue in the UK than in the US, as no single party advocates for traditional gender norms. Further, the UK has elected two women national leaders previously, whereas the US has yet to elect any, and has a history of higher women's representation in the legislature (IPU, 2022). It may be the case that British voters' greater familiarity with women politicians and leaders has led to less biased attitudes towards women politicians than in the US. While most experimental studies focus on only a single country, an important development for future work would be to run identical experiments in multiple countries simultaneously to understand how voter attitudes do, or do not, travel across countries and contexts. Perhaps most relevant would be to consider other Westminster-style parliamentary systems with a high proportion of women in the legislature and a his-

tory of women leaders, such as the New Zealand Parliament, or Western European democracies that far surpass both the UK and the US with respect to the presence and power of women in politics, such as the Swedish Riksdag or the Norwegian Storting. Given there is compelling evidence that different political systems and institutions, such as proportional representation, impact the descriptive representation of women (Matland and Studlar, 1996), we should also take seriously how these systematic differences may interact with how voters may or may not be biased towards women politicians.

The empirical analyses presented in this thesis concerns one central aspect of political behaviour: politicians' speechmaking behaviour and voters' engagement with this behaviour. The speeches that politicians deliver during plenary debates represent one of the richest sources of individual-level information on politicians' behaviour. Plenary debates are a fundamental part of the legislative and representational process, where politicians can engage in arguments over policy, articulate their views, represent their constituents, appeal to new voters, and hold the government to account. However, while these represent one of the key components of politicians' roles, speechmaking behaviour is only one component of politicians' behaviour. As discussed in chapter 2, communal and agentic stereotypes act as broad labels for many different features of politicians' styles, some of which extend beyond speechmaking to the wider way in which politicians' carry out their roles. The research I present makes an important contribution to existing work by producing new measures of several dimensions of politicians' styles at scale and over time, but there are several ways that future work can make progress on these efforts.

First, while much of style is captured in the content of what is said, facial expressions, vocal pitch, and body language also form central components of style. The use of text to study features of political behavioural styles in both a qualitative (Hargrave and Langengen, 2021; Haselmayer, Dingler and Jenny, 2021; Kathlene, 1994; Mattei, 1998) and quantitative (Gleason, 2020; Jones, 2016; Yu, 2013) manner now represents a common approach in the study of (gendered) legislative politics. However, a handful of recent studies have begun to develop measures for detecting non-verbal features of political style. Using a large quantity of audio data from the US Congress, Dietrich, Hayes and O'Brien (2019) introduce new measures to investigate the emotional intensity of legislators' speeches through studying small changes in vocal pitch. In doing so, they find that Congresswomen speak with greater emotional intensity when talking about issues related to women compared to their men colleagues or compared to when they discuss other topical areas. Studying videos from German federal election debates, Boussalis et al. (2021) use a variety of audio, video, and text approaches to measure candidate facial displays of emotion, vocal pitch, and sentiment of speech. By developing these approaches,

they study the emotional expressiveness of Angela Merkel compared to her opponents and, through leveraging live audience reactions, assess how voters react to Merkel's use of stereotype-incongruent emotions. They find that Merkel is rewarded for displays of happiness and punished for displays of anger.

These recent innovations by [Dietrich, Hayes and O'Brien \(2019\)](#) and [Boussalis et al. \(2021\)](#) demonstrate the opportunities researchers have to develop measures of political style such as emotion which, in theory, are detectable through vocal pitch, body language, or facial expressions. These approaches do, however, present hurdles with respect to data collection and availability. In the UK case, while resources such as Hansard and TheyWorkForYou allow researchers to gather large amounts of speech data spanning many decades with relative ease, video recordings of House of Commons debates are only available as far back as December 2007. Given my central substantive interest in chapters 3 and 4 both require the measurement of politicians' speeches over several years, it would not be possible to address these questions with a video analysis with the data available in the UK case. Nonetheless, there have yet to be any quantitative attempts to measure (gendered) political styles through audio or video analyses in the UK context, and this presents an important point of exploration for future work.

Second, while audio and video analyses enable researchers to delve into features of style beyond those measurable in the content of speech, these methods still do not allow researchers to measure stylistic traits that extend *beyond* speechmaking behaviour. In particular, a key dimension upon which women are said to differ from men is with respect to how *collaborative* they are. Collaboration captures the act of working with other legislators to craft policy, to network, to organise meetings, to exchange ideas and gather information, and to encourage colleagues to participate and speak out in meetings, floor debates, or committees. While my results suggest that women politicians are decreasingly using communal styles over time, the set of communal styles I focus on does not account for all features of style. Empirical work from contexts beyond the UK have assessed collaboration activities through interviews with politicians ([Barnes, 2016](#); [Tolleson-Rinehart, 2001](#)), bill co-sponsorship behaviour ([Barnes, 2016](#); [Holman, Merolla and Zechmeister, 2016](#); [Holman, Mahoney and Hurler, 2021](#); [Swift and Van der Molen, 2021](#)), and network analysis ([Wojcik and Mullenax, 2017](#)), and have found that women engage in these activities more than men. While particular forms of data, such as bill co-sponsorship, are not particularly meaningful in the UK case, future efforts could create a survey-based design in the spirit of [Wojcik and Mullenax \(2017\)](#) to address questions such as how often MPs work with one another, whether women are more likely to work with other women, or with politicians from parties beyond their own. Further, a valuable next step would also be to use an experimental design similar to the

one I present in chapter 5 to identify whether a collaborative political style is something that voters value. Such approaches would enable researchers to glean new insights into stereotype consistent styles beyond speechmaking alone, and to how voters engage with these styles.

Third, chapter 4 provides valuable insights into understanding how substantive representation shifts throughout politicians' careers, however this work does not enable me to identify how this translates into legislative voting behaviour and policy outcomes more broadly. An analysis of speechmaking behaviour provides me with valuable insights into the opinions, priorities, and representational efforts of politicians, but policy outcomes serving women's interests are ultimately more important for bringing about change. In the UK, it is challenging to gain insights into the roles of individual legislators in the policy-making process, because legislation is commonly proposed centrally by the government and MPs tend to vote in accordance with their party's positions. Nonetheless, understanding the extent to which this may translate into voting behaviour and policy outcomes presents an important challenge for future work.

Finally, in chapters 3 and 4 I emphasise the *dynamic* influence that stereotypes have on politicians' behaviour. In chapter 5, I show that voters do not seem to punish women for stereotype-incongruent behaviour, but I do not assess how these evaluations have shifted over time. Of course, as it is not possible to re-run the survey experiment at different points over the past 25 years, I cannot know whether UK voters in previous eras did hold women to stereotypical standards. However, future work should focus on multi-wave experiments to trace how voters' attitudes toward stereotype-(in)congruent politicians shift over time. Similarly, in chapter 3, I argue that the stylistic convergence of men and women in parliament reflects the declining influence of gender stereotypes in UK politics over time. The nature of the data prevents me from providing strong evidence regarding the *mechanisms* that led to these behavioural changes. For instance, voters may be less likely to endorse traditional gender stereotypes now than in the past, but I lack the data to assess whether there has been a concomitant decline in the sanctions that voters apply for gender-role-inconsistent behaviour. Future work on these issues should focus on collecting over-time survey data on voters' attitudes towards non-stereotypical behaviour by politicians, and on estimating *local* opinion on these issues which might then be linked to the behaviour of specific MPs in parliament.

For instance, I suggest that one important development relevant to the power of gender stereotypes is the increased prominence of women in parliament in recent years. However, the number of women politicians has increased monotonically over time in the UK, and so it is hard to distinguish the effect of this type of descriptive representation from other trends. Future work might use a design in

the spirit of [Karpowitz and Mendelberg \(2014\)](#), who randomly assign participants into political discussion groups with varying gender compositions. Do gender-stereotypical communication styles decline when more women are present in a group? Such efforts would provide both a more detailed insight into the changes in the extent to which UK voters uphold stereotypical standards and shed more light on the mechanisms that underpin the trends I document.

These concluding remarks highlight just a handful of ways, of which there are innumerable others, in which future research should consider furthering the study of gender, stereotypes, political style, and representation in legislative settings and beyond. Given the proliferation of new ways to gather data and innovation in techniques to measure elite political behaviour and voter attitudes, I hope the efforts in this thesis motivate others to delve deeper into the dynamic complexity of gendered norms, stereotypes, and how their influences may change over time and context.

# Appendix A

## Appendix A – No Longer Conforming to Stereotypes? Gender, Political Style, and Parliamentary Debate in the UK

### A.1 Word-embedding-based dictionaries

The word-embedding-based measurement strategy consists of several steps, which I describe in more detail in this section.

First, for each style I define a “seed” dictionary that represents the concept of interest. I use the following sources to construct the seed dictionaries:

1. **Affect** – Linguistic Inquiry and Word Count 2015 (Affect) ([Pennebaker et al., 2015](#))
2. **Fact** – Linguistic Inquiry and Word Count 2015 (Number and Quantitative) ([Pennebaker et al., 2015](#)) and all occurrences of any numeric figures
3. **Positive Emotion** – Regressive Imagery Dictionary (Emotions: Positive Affect) ([Martindale, 1990](#))
4. **Negative Emotion** – Regressive Imagery Dictionary (Emotions: Anxiety and Sadness) ([Martindale, 1990](#))
5. **Aggression** – A bespoke dictionary of words (see figure [A.1](#) below)
6. **Human Narrative** – A bespoke dictionary of words (see figure [A.2](#) below) and the 200 most common names of children born between 1970 and 2019

The final two seed dictionaries – which relate to aggression and human narrative – are my original constructions. These dictionaries were constructed by reading and watching debates from the House of Commons that are known to feature either aggression (for instance, Prime Minister’s Questions) or examples of human narrative (for instance, debates on mental health or social policy issues), and selecting words and phrases that I thought were likely to capture these concepts in a broader set of

parliamentary debates. I report the full lists of words that feature in these new seed dictionaries in figures A.1 and A.2

Figure A.1: Aggression seed dictionary

irritated ; stupid ; stubborn ; accusation ; accuse ; accusations; accusing ; anger ; angered ; annoyance ; annoyed ; attack ; insult ; insulting ; insulted ; betray; betrayed ; blame ; blamed ; blaming ; bitter; bitterly ; bitterness ; complain; complaining; confront ; confrontation; fibber; fabricator ; phony ; fibber ; sham ; deceived ; deceive ; disgrace; villain; good-for-nothing; hypocrite ; deception; steal ; needlessly; needless; criticise ; criticised ; criticising ; blackened ; fiddled; fiddle; problematic ; lawbreakers ; offenders; offend; unacceptable ; leech; phoney ; appalling ; incapable ; farcical ; absurd ; ludicrous; nonsense ; laughable ; nonsensical ; ridiculous; outraged ; hysterical ; adversarial ; aggressive ; shady ; stereotyping; unhelpful ; unnatural ; assaulted ; assault ; assaulting ; half-truths ; petty; humiliate ; humiliating ; confrontational; hate ; hatred ; furious ; hostile ; hostility ; nasty; obnoxious ; sleeze; sleezy ; inadequacy; faithless; neglectful ; neglect; neglected; wrong ; failure ; failures ; failed ; fail ; scapegoat ; cruel; cruelty ; demonise ; demonised ; tactic ; trick; trickery ; deceit ; dishonest ; deception; devious; deviousness; shenanigans ; fraudulence ; fraudulent ; fraud; swindling; archaic ; sly; slyness; silly; silliness ; scandal; scandalous ; slander ; slanderous ; libellous ; disreputable ; dishonourable ; shameful; atrocious ; gimmick ; immoral; ridicule; antagonistic ; antagonise ; ill-mannered; spiteful ; spite ; vindictive ; prejudice ; prejudices ; disregard ; arrogant ; arrogance ; embarrassment ; embarrass; embarrassing ; distasteful ; provoke; provoked ; petulant ; ignorance ; stupidity ; idiot ; idiotic ; annoying; dodgy ; untrue ; penny-pinching ; attacking ; ironic ; irony ; outrageous; hackery; crass; backchat; rude ; ill-judged ; ragbag; mess; hash ; fiasco; shambles ; shambolic ; farce; botch; botched ; blunder ; mischievous; mischief ; undermine ; straightjacket ; groan; abuse; chaos; chaotic ; dull; predictable ; negligent; grotesque; scapegoats; hypocrisy; bogus; counterproductive; betrayal; patronise ; patronising; reprehensible; fool; foolish; abysmal ; disgraceful; woeful; inferior ; sneaky ; scaremongering; scaremonger; coward; cowardly; ignorant; intolerant; unacceptable ; condemn; short-sighted; ashamed; falsehood; blackmail; clownery; debased; debase; hypocrisy; mislead; misleading; smokescreen; subterfuge; horrendous; despicable; deplorable

Figure A.2: Human narrative seed dictionary

example; constituent; person; someone; instance; surgery; case; told; illustrate; anecdote; experience; people; individual; cases; man; woman; mother; father; son; daughter; uncle; aunt; cousin; wife; husband; parent; child; say; said; support; discuss; speak; community; local; area; family; issues; remember; recall; married; resolve; authorities ; help; imagine; envisage; lives; sometimes; concerned; heard; circumstance; anyone; nobody; citizens; relationship; girl; boy; believe; listen; problem; inspire; many; comment; authority; conversation; worked; tell; thought; life; home; referred; situation; happened; everyone; concern; recognise; advice; advise; everyday; personal; letter; involve; nephew; niece; learn; local area; my constituents; previous job; tell me; told me; first hand; speaking as; own experience; for example; I recognise; I remember; help people; many years; see me; spoke with; their; them; talk; constituency ; constituents ; mum; dad; rhetoric; mr ; mrs ; know ; wrote ; write; ask ; call; dr; doctor; society; ordinary ; together ; dear; honest; visit; everybody; feel; view; public ; employer ; reflect; born; expect; anybody; responsibility ; youngster; heartbreaking; young; hopeless ; desperate; picture; chat; electorate; provide for; foster; colleague; represent ; neighbourhood; locality ; sympathy ; condolence; grief; bereavement ; trust; serve; communicate; testimony; motherhood; fatherhood; sensitive; remark; couple; brave; lifelong; proud; pride; facilities; quote; real; meet; met; childhood; reminisce ; nostalgia; recollect; hometown; lifetime; email; neighbour; partner; children; teenager; youth; contact; tale; scenario; bred; hard-working; year-old; friend; parent; parents; came; knew; recently; lady; gentleman; families



Second, a key component of the approach to measuring style are a set of word-embeddings, which I estimate from the full corpus of parliamentary speeches. Word-embedding models, which are of increasing use in political science (Rodriguez and Spirling, 2022), seek to describe any word in a corpus as a dense, real-valued vector of numbers. The construction of the word-embedding vectors, regardless of the specific algorithm used to estimate them, relies centrally on the distributional hypothesis: the idea that words which are used in similar contexts will have similar meanings. Here, a context refers to a window of words around a target word, and the embedding model allows me to *learn* the semantic meaning of each word directly from the use of the word in the corpus.

The main output of embedding models are the word-embeddings themselves. These are vectors that correspond to each unique word in the corpus. The dimensions of the embedding vectors capture different semantic “meanings” that can be used to provide structure to vocabulary. Crucially for my purposes, given this representation, the distances *between* word-vectors have been shown to effectively capture important semantic similarities between different words (Mikolov et al., 2013). I use this property to define the set of words that, *in the context of UK parliamentary debate*, are used in a semantically similar fashion to the seed words.

I follow the estimation procedure outlined in Pennington, Socher and Manning (2014) and estimate a word embedding,  $W$ , of length  $J = 150$  for each unique word in the corpus. I use a small “context” window size of 3 words either side of the target word to estimate the embeddings. This is consistent with my aim of capturing semantic (rather than topical) relations between words (Rodriguez and Spirling, 2022). I exclude all words that occur very rarely (fewer than 90 times overall), and all words that occur very frequently (in more than 90% of documents). I remove all stop-words, punctuation, and a bespoke list of parliamentary address terms such as “Honourable Friend” or “Home Secretary”. I collect the embeddings in a matrix,  $\theta$ , which I use to calculate the mean word-embedding vector for each of the seed dictionaries. The average word-embedding of the seed words represents the “location” of the dictionary in the vector-space defined by the embedding model, and allows me to calculate the relative semantic similarity of different words to the dictionary.

Third, I calculate the similarity between *every* word in the corpus and the mean dictionary word-vector using the cosine-similarity metric. Words closely related to the average semantic meaning of the seed words will have a high similarity score, and words that are less closely related will have a low similarity score. I then follow Zamani and Croft (2016) and apply the sigmoid function to the similarity scores, which transforms all similarity scores to the [0,1] interval and shrinks the scores of all but the most similar words to very close to zero. Where  $x_w^s$  is the cosine similarity between the word-

embedding for word  $w$  and the mean word-embedding of the seed dictionary for style  $s$ , the sigmoid transformation is given by:

$$Sim_w^s = \frac{1}{1 + e^{-a(x_w^s - c)}} \quad (A.1)$$

Here,  $a$  and  $c$  are free parameters which I set to be equal to 40 and .35, respectively, based on the results in [Zamani and Croft \(2016, 3\)](#).  $Sim_w^s$  gives the final score for each word for each style. Words closely related to the average semantic meaning of the seed words for a given dictionary will have a high  $Sim_w^s$ , and words that are less closely related will have a low  $Sim_w^s$ .

Finally, I use the word-level scores,  $Sim_w^s$ , to score each *sentence* in the corpus. As described in chapter 3, the score for a given sentence on a given dimension is:

$$Score_i^s = \frac{\sum_w^W Sim_w^s N_{wi}}{\sum_w^W N_{wi}} \quad (A.2)$$

where  $Sim_w^s$  is the similarity score defined above, and  $N_{wi}$  is the (weighted) number of times that word  $w$  appears in sentence  $i$ , where the weights are term-frequency inverse-document-frequency weights.<sup>1</sup>  $Score_i^s$  represents the fraction of words in sentence  $i$  that are relevant to dictionary  $s$ . When words with high scores for a given style appear frequently in a given sentence, the sentence will be scored as highly relevant to the style. The score for each *document* is then the weighted average of the relevant sentence level scores, where the weights are equal to the number of words in each sentence.

---

<sup>1</sup>TF-IDF weighting is used to down-weight very common words, and up-weight relatively rare words.

## A.2 Validation tests

As with all quantitative text analysis approaches, careful validation of the measures is essential (Grimmer and Stewart, 2013), and I provide two face validation checks in this section, in addition to further detail on the human validation task I present in chapter 3.

### *Word-level validation*

In table A.1, I examine the words that are associated with large  $Sim_w^s$  values for each of the styles. In particular, the table shows the top 30 words associated with each concept according to the word-embedding measure (*Top*), the words that are high-scoring based on the word-embedding measure, but which do not feature in the seed dictionaries (*Added*), and the words that are low-scoring on the word-embedding measure but which did feature in the seed dictionaries (*Removed*). The *Added* words are particularly important, as they represent words that are used in a similar context to the words in the seed dictionaries in the parliamentary setting, but which would be missed by traditional dictionary-based approaches.

The tables reveal that high-weight words (*Top*) generally correspond very closely to the style dimensions to which they relate. For instance, the top-loading words in the positive emotion dimension include “joy”, “delight”, “eager”, and “excitement”. Similarly, in the aggression dimension, top words include “disgraceful”, “shameful”, “outrageous”, and “scaremongering”. It is also encouraging that the top words in the face dimension are mostly numeric quantifiers, and the top human narrative words include “constituent”, “told”, “wrote”, “said”, and several words that indicate specific individuals (“son”, “father”, “wife”).

In addition, many words that are not included in the original seed dictionaries are nevertheless given high-weights via the word-embedding approach (*Added*). For example, the words “shocking”, “incompetence”, “pathetic”, and “deplore” do not appear in the aggression seed dictionary, but nevertheless receive high-weights for that style. That these words are consistent with intuitive notions of these broad stylistic categories, although not in the original dictionaries, highlights the fact that the word-embedding approach is successfully finding words that are semantically closely related to the key concepts of interest.

Similarly, the table also shows that some words included in the original seed dictionaries which are not semantically similar to the relevant concepts in the context of parliamentary debate are given low-weights by the word-embedding approach (*Removed*). For example, that “terrorism” is removed

from the negative emotion dictionary is encouraging, as within a parliamentary context the use of the word “terrorism” is likely to be from a reference to matters of policy rather than to an expression of emotion.

Overall, the words in table A.1 suggest that the word-embedding model is a) accurately associating sensible words with the stylistic concepts; and b) capturing language use that is representative of a given style, even when those words are not included in the seed dictionaries, and so would be missed by traditional dictionary approaches.

<b>Affect</b>			<b>Positive Emotion</b>			<b>Negative Emotion</b>		
<i>Top</i>	<i>Added</i>	<i>Removed</i>	<i>Top</i>	<i>Added</i>	<i>Removed</i>	<i>Top</i>	<i>Added</i>	<i>Removed</i>
feel	feel	award-winning	joy	eager	gladstone	upset	upset	painstaking
really	really	admiral	delight	anticipation	reliefs	suffering	terrible	painting
sometimes	sometimes	securities	eager	pity	satisfied	terrible	hurt	alarms
afraid	undoubtedly	super	enjoyable	liked	relieve	distressing	deeply	paint
fear	frankly	destroyers	happy	hear	relieving	hurt	unfortunate	paints
undoubtedly	always	approvals	excitement	appreciated	satisfactorily	distress	angry	terrific
frankly	think	festival	enjoying	amazed	satisfy	frightening	felt	disappointingly
always	nevertheless	dwellings	cheer	wonderful	relief	unhappy	feeling	terrorists
think	often	engagements	celebration	sadness	gay	worry	caused	avoidance
nevertheless	genuinely	championships	delighted	love	grind	deeply	horrendous	cowardly
often	believe	championship	relieved	doubtless	satisfies	dreadful	appalling	grievance
genuinely	seem	approving	celebrate	birthday	satisfactory	unfortunate	shocked	hopelessly
believe	felt	shakespeare	amused	horrified	entertainment	worried	frustrating	lone
certainly	however	challenger	anticipation	always	grinding	suffer	compounded	miserably
seem	indeed	treasurer	fun	praise	amusement	anxiety	anger	terrorism
felt	feeling	pesticides	entertaining	informative	laughed	despair	frustration	alarmingly
however	perhaps	approved	pity	fascinating	laughs	fear	sometimes	grievances
indeed	obviously	risk-based	enjoyed	churlish	satisfaction	frightened	horrible	alarmist
feeling	something	harmonise	liked	pleased	gladly	angry	experiencing	painting
perhaps	say	flexibilities	hear	admire	cheers	suffered	feel	timid
obviously	probably	energy-intensive	appreciated	christmas	satisfying	sad	frustrated	terrorist
worry	find	relaxing	enjoy	look_forward	laughing	felt	appalled	discouraged
something	deeply	laughs	excited	afternoon	entertain	feeling	shocking	shy
say	nothing	approve	amazed	lovely	laughable	tragic	understandably	discouraging
probably	people	festivals	wonderful	fascinated	rejoice	caused	disturbed	avoids
find	thing	exhaustive	sadness	spirit	cheered	horror	unpleasant	lamentable
deeply	suspect	glamorgan	love	compliment	enthusiastically	horrendous	embarrassing	pitiful
nothing	somehow	approves	celebrating	astonished	enjoyment	appalling	terribly	discourage
people	quite	resignations	doubtless	coincidence	celebrates	shocked	frankly	sufferers
thing	much	praises	glad	sincerely	joke	frustrating	imagine	painfully

<b>Aggression</b>			<b>Fact</b>			<b>Human Narrative</b>		
<i>Top</i>	<i>Added</i>	<i>Removed</i>	<i>Top</i>	<i>Added</i>	<i>Removed</i>	<i>Top</i>	<i>Added</i>	<i>Removed</i>
disgraceful	utterly	inferior	half	nearly	sixthly	constituent	like	poppy
shameful	cynical	offenders	five	year	sevenoaks	told	called	bred
outrageous	frankly	assaulted	four	whereas	doubly	know	whose	amber
scaremongering	embarrassing	annoyance	nearly	years	infinitely	wrote	went	florence
utterly	incompetence	fiddle	three	£	ooost-century	like	think	georgia
cynical	misguided	fiddled	ooo	months	double-dip	called	indeed	anecdote
frankly	irresponsible	steal	six	just	infinite	said	says	hopeless
scandalous	pathetic	assault	year	days	oooth-century	whose	also	recollect
dishonest	dreadful	offend	whereas	weeks	scarce	constituents	just	alice
embarrassing	bizarre	furious	years	moreover	bunch	father	others	aunt
absurd	complacency	fail	seven	past	seven-day	mr	asked	eve
ridiculous	illogical	deceived	quarter	compared	groupings	son	saying	skye
ludicrous	incompetent	predictable	two	almost	fifthly	tell	week	albert
deplorable	shocking	dodgy	eight	yet	samples	went	see	chat
incompetence	reckless	fool	£	now	grouped	met	wanted	spencer
misguided	disingenuous	problematic	million	next	ooo-page	remember	perhaps	kate
irresponsible	complacent	bitterness	months	spend	equalities	think	former	mohammed
pathetic	unfortunate	fiasco	billion	ago	equalise	indeed	described	rhetoric
appalling	downright	neglected	just	thirds	ooog	says	obviously	ashton
dreadful	deliberate	betray	average	figure	group's	dr	one	tale
nonsense	wicked	cruelty	least	addition	ooond	wife	ooo-year-old	roman
bizarre	unjust	confrontational	days	roughly	sixth-form	david	mine	inspire
complacency	deplore	deceive	weeks	week	oob	say	knows	youngster
ashamed	unacceptable	archaic	moreover	furthermore	equalisation	also	yesterday	nicola
illogical	plainly	blackmail	past	number	ooo-to-ooo	just	unfortunately	locality
arrogant	horrible	embarrass	compared	within	six-week	family	friends	everyday
incompetent	manifestly	mischief	third	times	triple	others	looked	sensitive
arrogance	callous	smokescreen	almost	april	four-year-old	woman	aware	jamie
accusation	somehow	needlessly	one	probably	grouping	asked	although	carter
shocking	muddle	adversarial	yet	equivalent	oord	man	now	scenario

Table A.1: Word-level validation

*Sentence-level validation*

Tables A.2 and A.3 assess the face validity of the approach by showing the 10 highest scoring *sentences* for each style, according to the  $Score_i^s$  measure described in equation A.2. For all styles, the sentences clearly reflect the conceptual definitions I outline in table 3.1 in chapter 3. For instance, the “fact” category is dominated by statements using numerical language, and the “human narrative” category has many examples of MPs referring to the experiences of specific individuals. This again suggests that the measurement strategy plausibly captures my stylistic dimensions of interest.

Table A.2: Top sentences for Affect, Positive Emotion, and Human Narrative

Affect	Positive Emotion	Human Narrative
Others eventually got jobs, although usually far less rewarding, far less secure and far less well paid.	As always, it is an enormous pleasure to follow the hon Member for Bootle , whose speeches are always entertaining and occasionally informative.	Moreover, what happens when an elderly brother and sister live together, or an elderly mother lives with her elderly son?
In others, everyone seems a little depressed - perhaps not greatly upset but a little depressed none the less.	It is always a pleasure to listen to Members' maiden speeches, and I enjoyed his as well.	Last week a friend of mine who works with elderly residents in Ogmores visited four elderly residents in one day.
Some of us believe that the legislation is profoundly unacceptable, profoundly wrong and profoundly damaging to our country.	I am always excited and in a state of eager anticipation to hear what the right hon Gentleman has to say on everything.	Anyone whose wife or partner had a child 20 years ago will remember that the woman spent a week to two weeks in hospital.
We also need to stop trying to blame someone every time something bad happens: sometimes bad things happen and they are no one's fault.	I begin this afternoon by wishing the Secretary of State a very happy birthday - I sincerely hope that it improves from here on.	However his father David suffered a stroke 13 years ago since when his mother Sarah has had to care for both son and husband.
Such serious problems have left many facing uncertainty, which can cause severe stress to people who already face incredibly challenging circumstances.	I join hon Members across the House in wishing a happy Pride to all those celebrating London Pride this weekend.	I speak as someone whose father served in the Metropolitan police for 25 years and whose younger brother is a serving Metropolitan police officer.
Many mentally ill people face sad and painful lives with great courage - more courage than the rest of us may have.	I hope that I have the pleasure of listening to his own speech today, because I enjoy his speeches immensely.	American civilians took leave once every six months; British diplomats took leave every six weeks, for two weeks.
Is it any wonder that mentally ill people desperate for help just get lost, sometimes with tragic consequences?	I also congratulate my hon Friend the Member for Blackpool, North on his most amusing, entertaining and sincere maiden speech.	I have also discovered that a person called Mr Richard Shires subsequently became a paid constable in West Yorkshire police and continues to serve to this day.
All of us are aware that the Labour party has trouble understanding aspiration and even more trouble in rewarding aspiration.	Today's debate has been extremely lively, interesting and, at times, amusing and much good wit and humour have made it a delight.	On 13 March 1942, in New End hospital, the older brother that I never knew, James John Dromey, died at three days old.
People understandably already feel fraught and upset - they are in a situation that they never anticipated, and feel vulnerable and sometimes deeply hurt and angry.	I had a great surprise last Christmas when I received both a birthday card and a Christmas card from John and his family.	On Monday this week, another south Birmingham MP and I met South Birmingham primary care trust to talk about the situation in south Birmingham.
But neither can anyone underestimate the anger and sadness among people that things should ever have been allowed to get into this position.	It was wonderful to hear the shadow Chancellor - it is always wonderful to hear the shadow Chancellor in his marvellous speeches - explaining how cross-party he was.	Yes, another day, another Home Office statement and, sadly, yet another similar response from the shadow Home Secretary.



Table A.3: Top sentences for Aggression, Fact, and Negative Emotion

Aggression	Fact	Negative Emotion
I found the attitude of the Conservatives' motion not only hypocritical and incoherent, but profoundly cynical and dishonest.	None the less, social security on average now costs every working person nearly £15 every working day.	People understandably already feel fraught and upset - they are in a situation that they never anticipated, and feel vulnerable and sometimes deeply hurt and angry.
That statement is as barbaric as it is downright stupid; it is nothing more than an ignorant, cruel and deliberate misconception to hide behind.	The growth rate figures are substantially different from the growth rate figures produced in the Budget just four months ago.	However, the indignity, discomfort and inconvenience caused to Brian during this episode understandably left him feeling demoralised and, in his words, depressed.
It is grossly irresponsible and, I am afraid, profoundly and disturbingly misleading, and even ignorant, to go around doing that.	I have primary schools receiving less than £3,500 per pupil and secondary schools receiving less than £4,600 per pupil.	This is deeply worrying for families living in those blocks, and is causing huge anxiety, fear and insecurity.
There is something horrible, vindictive and cowardly about the Government's intolerant and ignorant attack on a small minority".	The maximum figure for those costs was \$91 billion, although the real extra costs amounted to \$26 billion.	Such serious problems have left many facing uncertainty, which can cause severe stress to people who already face incredibly challenging circumstances.
They should not be all about blaming people, because blaming individuals for errors and mistakes is unhelpful and counter-productive.	Recent figures show the current account deficit running at the much lower level of £0.5 billion per month.	It can cause misery and pain for individuals and their families through serious disease or, worse, death.
Of course the situation in Zimbabwe is disgraceful and we condemn utterly the barbaric attacks on farmers, which are totally unacceptable.	We now spend nearly £11 billion extra each year on pensioners, and almost half that additional spending goes to the poorest third.	In addition to suffering horrendous physical injuries, enormous physical stress and emotional trauma, they had enormous financial stress.
Some of us believe that the legislation is profoundly unacceptable, profoundly wrong and profoundly damaging to our country.	The five Conservative speakers took three hours, five minutes; the six Labour speakers took one hour forty-five minutes.	Children described the extreme distress they experienced: losing weight, having nightmares, suffering from insomnia, crying frequently and becoming deeply unhappy.
Worse even than the failure publicly to criticise and condemn has been the United Kingdom Government's tendency almost to excuse.	They would produce sentences of seven months, six days or nine months, six days and various split months and split days.	If people feel isolated, depressed, lonely, jobless and skill-less, they will feel worse in hospital.
To claim that the financial crisis was somehow caused by the Labour party's mismanagement is complete and utter nonsense.	Working in early years or later years care in private services means earning minimum wage or minimum wage plus.	To the families we say: we are deeply sorry for your loss and deeply sorry for the pain you have suffered.
If that happens because of an arrogant and incompetent subordinate should not that arrogant and incompetent subordinate be fired?	Approximately 100 people per 1,000 currently receive disability living allowance, compared with 50 people per 1,000 in Britain.	Their anger is the anger of pain, the anger of discrimination, and the anger of lack of understanding, as well as the anger of frustration.

### A.3 Controlling for individual-level covariates

In this section, I show results of the alternative specification for the dynamic hierarchical model described in chapter 3, in which I expand the model at the second level by including a vector of individual-level covariates,  $X_{j,t}^k$ :

$$\alpha_{j,t} \sim N(\mu_{0,t} + \mu_{1,t} \text{Woman}_j + \sum_{k=1}^k \lambda_k X_{j,t}^k, \sigma_\alpha) \quad (\text{A.3})$$

where  $X_{j,t}^k$  includes:

- Party (categorical: Conservative; Labour; Liberal Democrat; Other)
- Government or opposition party status (binary)
- Government or opposition frontbench position (binary)
- Committee chair (binary)
- MP age (in years, continuous)
- Margin of victory in prior election (percentage points, continuous)
- University degree (binary)
- Prior occupation (categorical: manual; professional; political; business; other)

I transform the two continuous predictors such that they have mean zero, and standard deviation one. I present the results for the main quantities of interest ( $\mu_{1,t}$ ) estimated from this model in figure A.3.

The figure shows that, in general, I recover very similar patterns of gender differences in style use over time when controlling for individual-level covariates. For human narrative, affect, positive emotion, negative emotion, fact, and aggression the trajectories of the gender differences over time are very similar to those presented in figure 3.2 in chapter 3. The largest differences are for complexity and repetition, where the pattern of convergence between men and women is somewhat attenuated in the estimates from the alternative specification. For complexity in particular, the large shift in the gender difference that I observe between 2008 and 2013 is confounded by some of the individual-level covariates, as the gender difference is largely constant (and indistinguishable from zero) for the entire time period once I control for these other factors. Nevertheless, overall, these results suggest that while other MP-level characteristics clearly account for some variation in style use, my central finding – that the debating styles of men and women MPs have diverged from gender-based stereotypes over

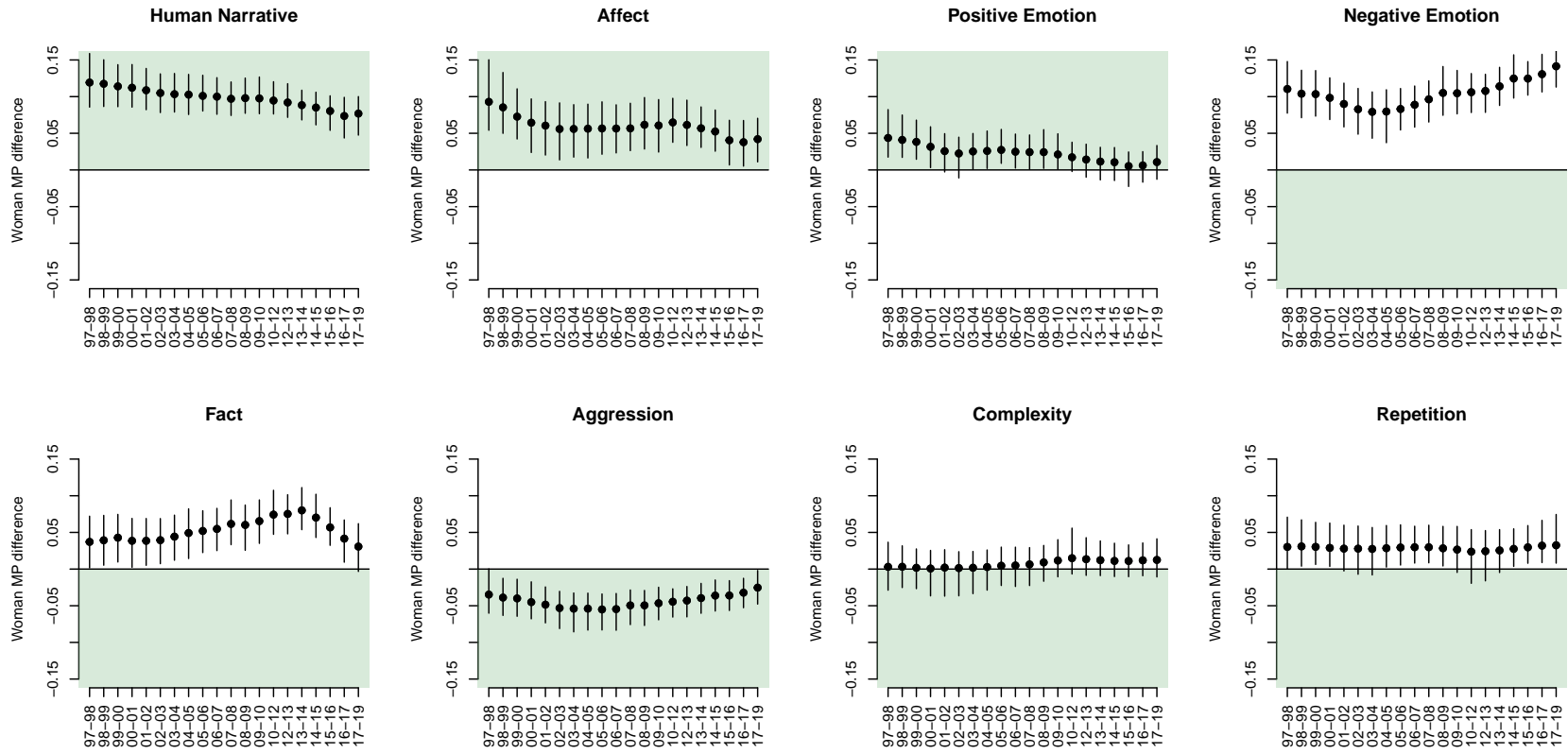


Figure A.3: Gender differences in style over time controlling for individual-level confounders

time – is not affected by these estimates.

Figure A.4 presents the estimates for each of the individual-level covariates for each style. Although these are not my primary quantities of interest, there are several patterns that are of substantive interest. First, I find, consistent with other work (Proksch et al., 2019), that MPs from government parties use significantly less negative and more positive language than MPs from opposition parties. Government MPs are also less aggressive and tend to rely more on human narrative and less on fact-based arguments than their opposition counterparts. Second, compared with backbench MPs, politicians in leadership positions are less likely to use human narrative, more likely to make fact-based arguments, use substantially less emotive language, and are more repetitious in their speeches. I also see some evidence of partisan differences. Compared to Conservative Party MPs, Labour MPs use more human narrative, more factual language, and are somewhat less complex in their speeches. Liberal Democrat MPs, by contrast, make less use of human narrative, more use of fact, and are substantially less aggressive than Conservative MPs. There are also interesting patterns in speech styles according to the education and occupation variables. For instance, university-educated MPs tend to make less use of human narrative, and less use of negative emotional language, but deliver speeches that are more complex and more repetitious than their non-university educated counterparts. With regard to prior employment, MPs from manual occupations do appear to have distinct speechmaking styles, as they employ more human narrative, and less aggressive and repetitive language than MPs from other employment backgrounds. Overall, it is clear that there are many factors that influence the political styles that MPs adopt and, while these are not directly relevant to the substantive questions in my study, I think that these findings may be profitably investigated in future work.

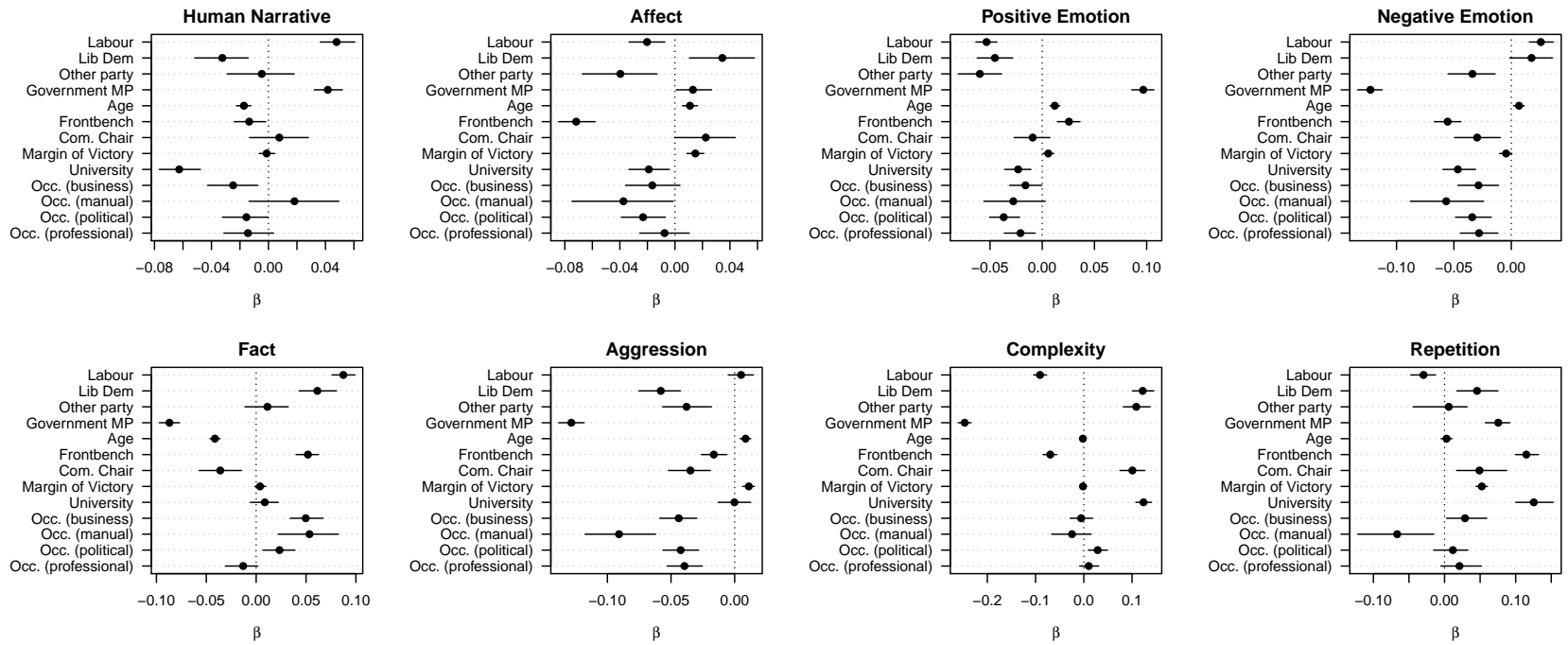


Figure A.4: Individual-level covariate effects

## A.4 Style use and debate-type

My model accounts for aggregate differences in style use across debates via the  $\delta_d$  random-effects described in equation 3.2 in chapter 3. The inclusion of these parameters means that gender differences in style use cannot be attributed to men and women participating in systematically different types of debates, as the gender effects I estimate are based on within-debate variation in the style outcomes. However, it is possible that the magnitude of gender differences nevertheless varies across debates of different types. I investigate this possibility here. Specifically, I separate the debates in the data into common types that occur regularly in the UK House of Commons (for more detail, see [Blumenau and Damiani, 2021](#)):

1. **All:** all debates in the dataset.
2. **Ministerial Question Time:** the routine questioning of Ministers, occurs four times a week.
3. **Prime Minister's Question Time:** the Prime Minister answers questions from the Leader of the Opposition, opposition members and government backbenchers, occurs once a week.
4. **Procedural debates:** a compound category that includes debates that are not substantive in nature, but deal with matters of parliamentary procedure or scheduling. For example, Business of the House or Points of Order.
5. **Legislation:** debates on legislation, includes all stages of the process that occur in the House of Commons chamber, such as second and third reading.
6. **Opposition Days and Backbench Business:** this includes business for debate that is placed on the parliamentary agenda by opposition members or backbenchers.
7. **Other:** all other forms of debate that are not captured by the above categories.

This categorisation captures important substantive differences between different types of debates in the House of Commons, some of which have been shown to be predictive of MPs' style in previous work ([Osnabrügge, Hobolt and Rodon, 2021](#)).

I run a series of OLS models for each of the style-based outcomes, where my main explanatory variable of interest is the gender of the MP, and where I also control for party, age, years in parliament, margin of victory in the previous election, degree education, previous occupation, and whether the MP was a) a member of the cabinet, b) a member of the shadow cabinet membership, c) a government minister, d) a shadow minister, or e) a committee chair. For each outcome, I subset the data to only debates of a certain type, estimate the model, and record the coefficient on the gender variable at each iteration. Figure A.5 shows, for each style, the gender differences in the different debate types.

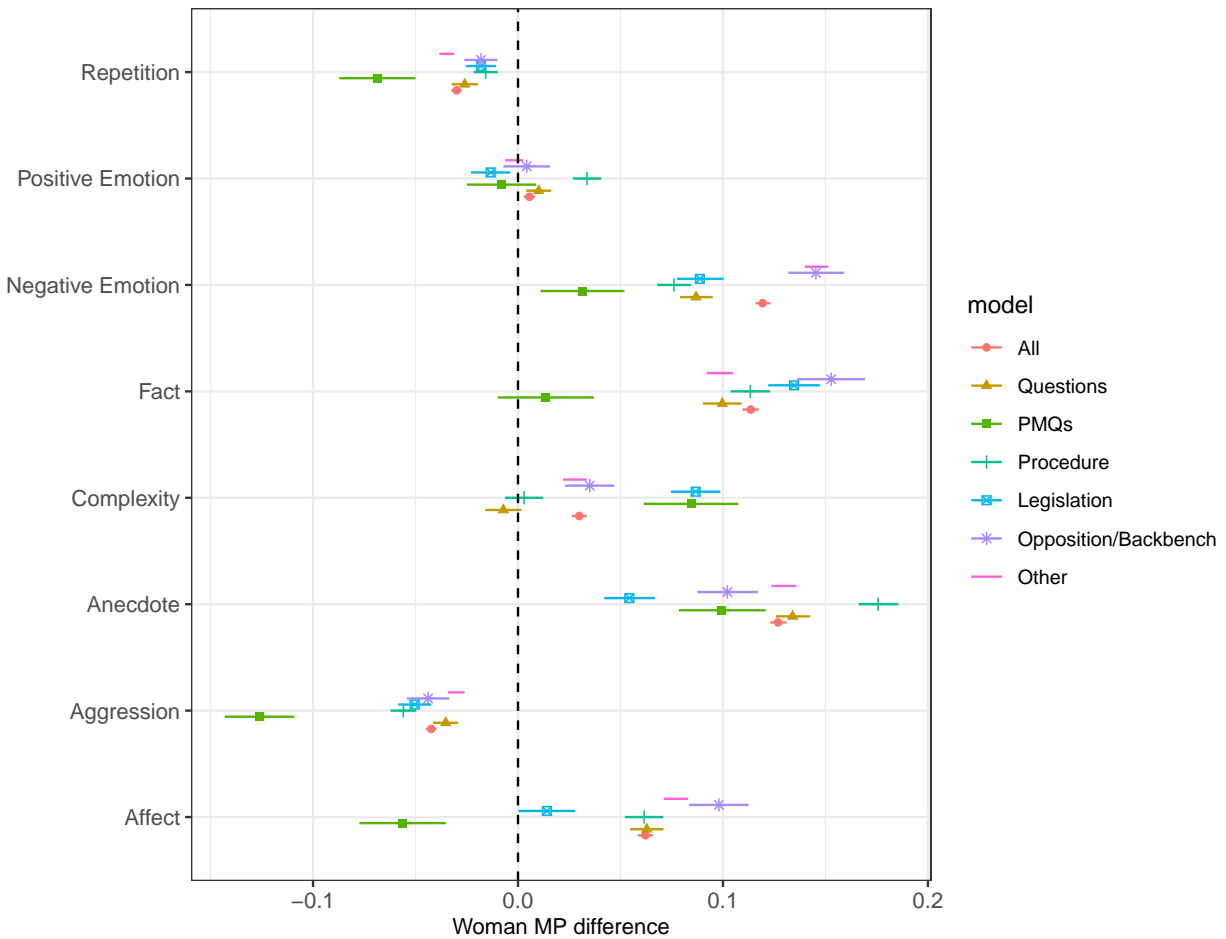


Figure A.5: Debate type models

The analysis reveals that the magnitude of average gender differences are relatively constant across the debate types. In the debate types I identify, Prime Minister's Questions seems to be the only type of debate that significantly effects the gender coefficients. I see that, relative to the model which pools across all debates, the magnitude of gender differences is increased for repetition, aggression, and affect; decreased for negative emotion; and reduces gender differences in fact to statistically indistinguishable from zero. Overall, however, while there is some variation in the magnitude of gender differences across debate types, these differences are for the most part very small.

In figure A.6, I show additional descriptive information on the average level of each style in speeches used across the different debate types. The patterns in style use across debates generally conform with standard intuitions. For instance, the figure shows that both Question Time and Prime Ministers Questions (PMQ) debates are substantially less positive than debates on legislation, which

is consistent with the idea that these settings are used by the opposition parties to interrogate – and often castigate – the government on issues of the day. Similarly, both PMQ debates and debates initiated by the Opposition parties in parliament are more aggressive than other debates, which again follows the intuition that these debates are mainly used as a vehicle for criticising government policy. In general, these descriptive figures bolster the results from the validation exercises above, as they imply that the measures accurately capture expected differences in speech style across different types of parliamentary debate.



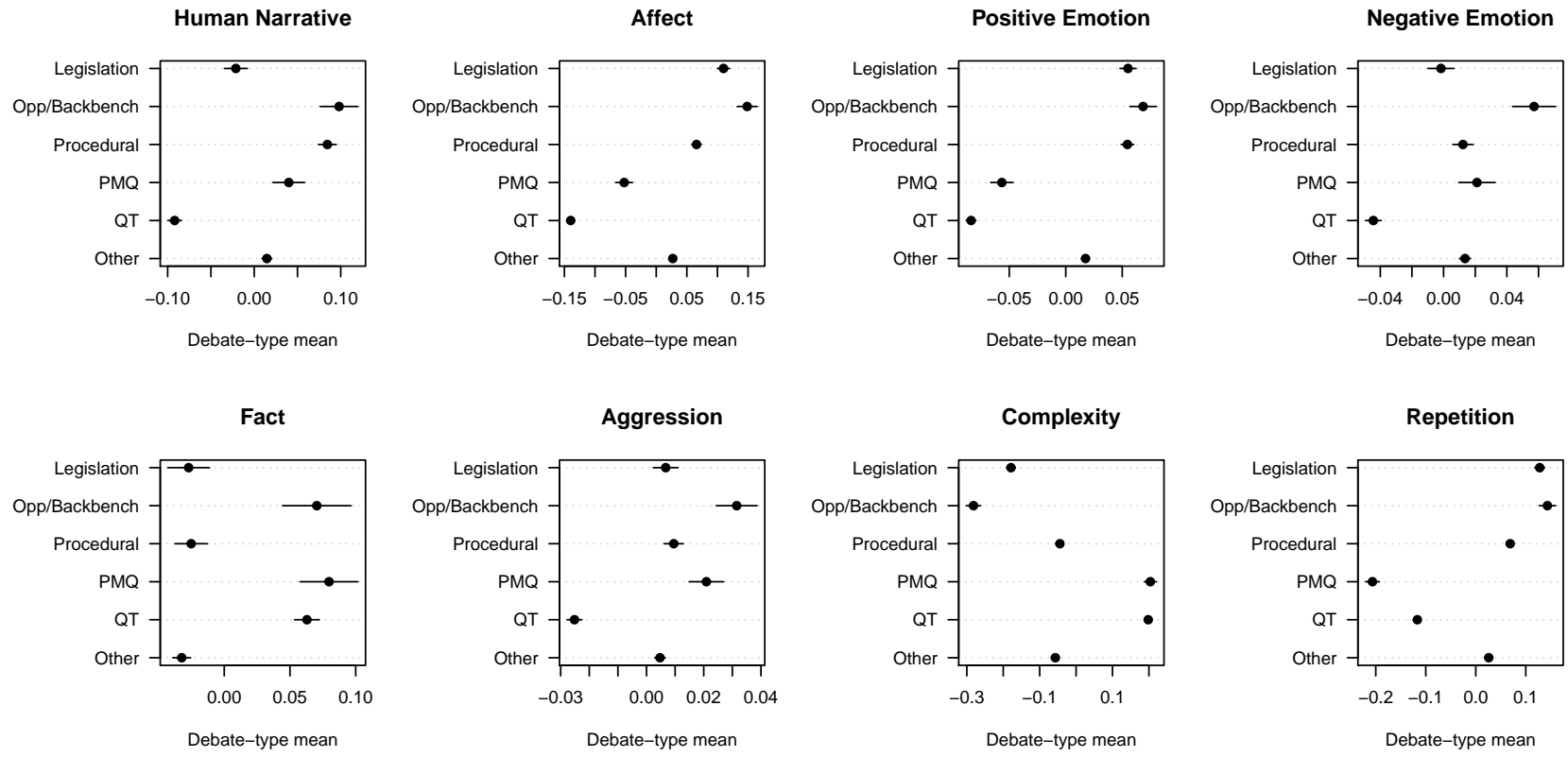


Figure A.6: Style type average by debate type

## A.5 Within-MP and replacement effects

Does gender explain less variation in aggregate style use over time because of a gradual convergence in styles of men and women MPs throughout their careers in parliament? Or do gender gaps decrease because the men and women entering parliament over time are systematically different from those leaving parliament? Which of these two explanations – which I refer to as “within-MP” and “replacement” effects – is responsible for the aggregate patterns I document in chapter 3? The modelling approach allows me to decompose the evolving gender differences that I report in chapter 3 into these two mechanisms of change.

Given the model described by equations 3.2 and 3.3 in chapter 3, I can decompose the shifting patterns of gendered style use into those changes that stem from within-MP change over time, and those that come from replacement. My goal is to specify a decomposition of  $\mu_{0,t} - \mu_{0,t-1}$ , which is the change in average style use for men between parliamentary session  $t$  and session  $t - 1$  (I can then provide an equivalent approach for women MPs). I begin by distinguishing between three types of MP, which I label as “remainers”, “joiners”, and “leavers”:

- $J_m^R$  is the set of men who appear in both session  $t$  and  $t - 1$  (Remainers)
- $J_m^J$  is the set of men who appear in  $t$  but not in  $t - 1$  (Joiners)
- $J_m^L$  is the set who appear in  $t - 1$  and not in  $t$  (Leavers)

I also require the fraction of men who are “remainers” in  $t$  and  $t - 1$ :

- $\pi_t^R$  is the fraction of men in  $t$  who also served in  $t - 1$
- $\pi_{t-1}^R$  is the fraction of men in  $t - 1$  who also served in  $t$

Note that the proportion of men who are “remainers” in  $t$  may be different from the proportion in  $t - 1$ , because some men who leave parliament in  $t - 1$  will be replaced by women in  $t$  (and vice versa).

Given these definitions, I can write the mean style use for men in each period as a function of the MP-period effects ( $\alpha_{j,t}$ ):

$$\mu_{0,t-1}^m = \underbrace{\pi_{t-1}^R \frac{1}{|J_m^R|} \sum_{j \in J_m^R} \alpha_{j,t-1}}_{\text{Remaining MPs}} + \underbrace{(1 - \pi_{t-1}^R) \frac{1}{|J_m^L|} \sum_{j \in J_m^L} \alpha_{j,t-1}}_{\text{Leaving MPs}} \quad (\text{A.4})$$

$$\mu_{0,t}^m = \underbrace{\pi_t^R \frac{1}{|J_m^R|} \sum_{j \in J_m^R} \alpha_{j,t}}_{\text{Remaining MPs}} + \underbrace{(1 - \pi_t^R) \frac{1}{|J_m^J|} \sum_{j \in J_m^J} \alpha_{j,t}}_{\text{Joining MPs}} \quad (\text{A.5})$$

Here,  $\mu_{0,t-1}$  is a weighted average of the finite-sample average of the “remainers” and “leavers” in  $t - 1$ , where the weights are given by the relative proportion of those groups in that parliamentary session.  $\mu_{0,t}$  is constituted from the equivalent averages for “remainers” and “joiners” in time period  $t$ , again weighted by the size of those two groups in  $t$ .

Taking the difference between A.4 and A.5 and rearranging reveals an additive decomposition which separates the two effects of interest:

$$\begin{aligned} \mu_{0,t}^m - \mu_{0,t-1}^m &= \underbrace{\pi_t^R \frac{1}{|J_m^R|} \sum_{j \in J_m^R} \alpha_{j,t} - \pi_{t-1}^R \frac{1}{|J_m^R|} \sum_{j \in J_m^R} \alpha_{j,t-1}}_{\text{“Within-MP” effect } (S_m)} + \\ &\quad \underbrace{(1 - \pi_t^R) \frac{1}{|J_m^J|} \sum_{j \in J_m^J} \alpha_{j,t} - (1 - \pi_{t-1}^R) \frac{1}{|J_m^L|} \sum_{j \in J_m^L} \alpha_{j,t-1}}_{\text{“Replacement” effect } (R_m)} \end{aligned} \quad (\text{A.6})$$

I denote the within-MP effect for men as  $W_m$  and the replacement effect as  $R_m$ . I can also, of course, define the same quantities for women MPs, and therefore can describe the changing gender difference in terms of replacement and socialisation effects:

$$(\mu_{0,t}^w - \mu_{0,t}^m) - (\mu_{0,t-1}^w - \mu_{0,t-1}^m) = \underbrace{(W_w - W_m)}_{\text{“Within-MP” difference}} - \underbrace{(R_w - R_m)}_{\text{“Replacement” difference}} \quad (\text{A.7})$$

Turning to the results, I plot these quantities in the left (for men) and centre (for women) panels of figure A.7. The x-axis describes the average direction and magnitude of changes between parliamentary sessions for each style for men and women, respectively. The right-hand panel reports *the difference* in the effects for women and men. In each panel, hollow points show changes that occur because of replacement, and solid points show changes that occur due to within-MP shifts.

I use these plots to understand whether replacement or within-MP change is a stronger determinant of the aggregate shifts I observe. Overall, neither differential replacement nor within-MP change alone explain the convergence that I document across multiple different styles in chapter 3, though there is some evidence that replacement is more important as a mechanism for explaining the changing gender dynamics I observe for the agentic styles while within-MP change is somewhat more important for explaining change for more communal styles.

For example, figure 3.3 in chapter 3 shows that women are much more likely than men to use negative emotion in their speeches in later years, but only somewhat more likely in the earlier years.

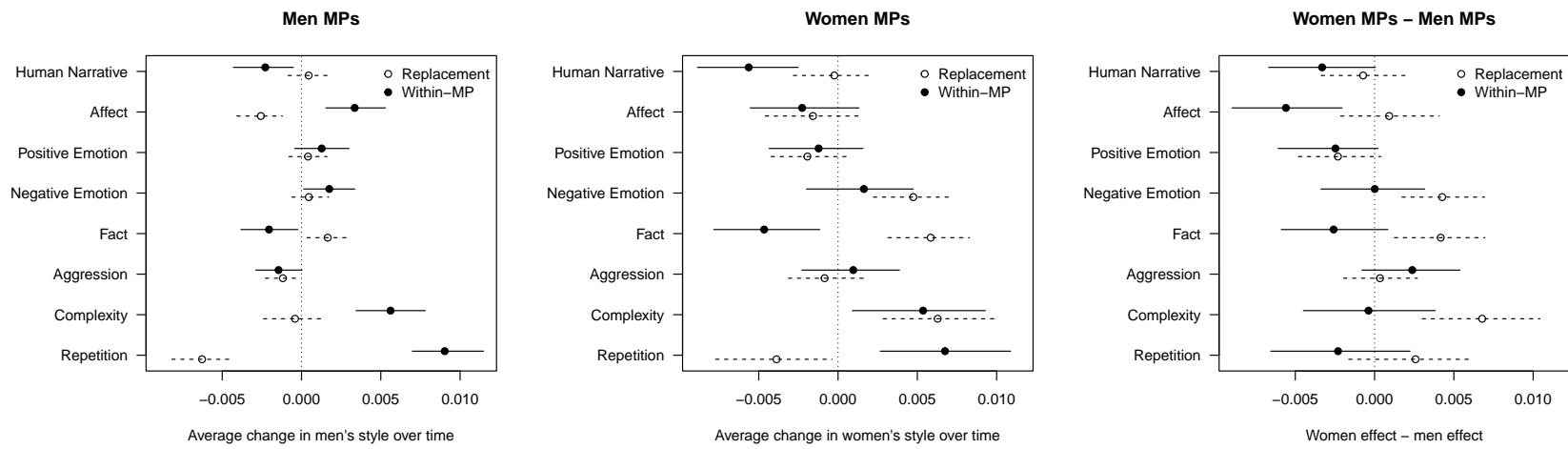


Figure A.7: Within-MP and replacement over time change by gender

The middle panel of figure A.7 shows that the replacement effect for women for negative emotion is positive (the hollow point for negative emotion is greater than zero), which implies that newly elected women are more negative than the women leaving parliament, on average. However, the left panel of figure A.7 suggests that this is *not* true for men: men joining parliament use negative language at the same rate on average as men leaving parliament (the hollow point for negative emotion is close to zero). Consequently, the right panel suggests a (positive) differential replacement effect for negative emotion. Note that the difference between men and women *within-MP* effects is close to zero for negative emotion. This suggests that the divergence between men and women that I note at the aggregate level is almost entirely driven by differential replacement between men and women MPs, rather than existing MPs becoming more alike in their behaviour over time.

The right-hand panel of A.7 indicates that, beyond negative emotion, replacement effects also account for a greater share of the aggregate change in gender differences for factual language and complexity. For both, while the women entering parliament are significantly more likely to use these styles than the women leaving, newly elected men employ these styles at broadly similar rates as the men that they replace. By contrast, both men *and* women MPs are less likely to use factual language as their careers in parliament progress, and the speeches of both men *and* women become more complex the longer they spend in parliament. Consequently, the large aggregate shifts that I observe for these styles are largely driven by the fact that the women newly elected to parliament adopted a legislative debating style that was more factual, complex, and negative than the women they replaced.

For other styles, I see that within-MP change accounts for a greater share of the variation in gender differences. For instance, the gradually decreasing gap in the use of human narrative in figure 3.3 in chapter 3 is mostly attributable to women MPs using this style less the longer that they stay in parliament, but the decreasing use of human narrative for men is much smaller. Similarly, on average women employ less positive emotion over time, whereas the positive language use of men remains relatively constant. Conversely, within-MP change in affect for men is positive, implying that men become more emotional overall in their speeches over time, but there is very little average within-MP change in affect for women. These results imply that, for these styles, the convergence that I see in the main analysis in chapter 3 is driven by the different stylistic trajectories than men and women MPs appear to follow throughout their tenure in parliament.

## A.6 Topic-based confounding

I present evidence of convergence between men and women with respect to several debating styles over time. One potential concern for the interpretation of these results is that the parliamentary agenda is not fixed, and changes to the set of issues under discussion may result in convergence between men and women even in the absence of behaviour change.

Consider, for instance, a style like human narrative, where I observe a large convergence between men and women over time. Women are significantly more likely to use human narrative in their parliamentary speeches at the beginning of the time period than they are at the end. If, however, women are more likely to use human narrative than men in certain *topics*, and those topics become less prevalent over time, then the convergence I document might in fact be attributable to changes to the parliamentary agenda. For changes in topic prevalence to be responsible for convergence, it would have to be the case that the topics on which I observe women using *more* human narrative than men are becoming *less* prevalent, or that the topics on which I see women using *less* human narrative than men are become *more* prevalent over time. For example, perhaps women use more human narrative than men when discussing education policy, and education policy is more frequently discussed in the early period in my data than the later period in the data. If this were true, then the results might be subject to topical confounding, as changes in topical prevalence over time would account for the aggregate changes I observe in the main analysis.

To address this concern, in this section, I use statistical topic models to evaluate whether topics on which I observe notable stylistic differences between men and women become more or less prevalent over time. I begin by estimating a correlated topic model (Blei and Lafferty, 2006) (CTM) for all speeches in the data. The CTM is an unsupervised learning approach which assumes that the frequency with which words co-occur within different speeches provides information about the topics that feature in those speeches. As with other topic models, the CTM requires the analyst to choose the number of topics,  $K$ . Given that the results might be sensitive to this choice, I choose to present results from a series of models, where I vary the number of topics:  $K \in 10, 20, \dots, 80$ . I implement the CTM as the null form of the Structural Topic Model, which I implement in R (Roberts, Stewart and Tingley, 2014).

The key output of the topic model is  $\theta$ , a  $N * D$  matrix of topic proportions that measures the degree to which each speech ( $i$ ) in the data features each of the estimated topics ( $d$ ).  $\theta_{i,k}$  therefore gives the proportion of speech  $i$  devoted to topic  $d$ . With these topics in hand, I then evaluate – for each of the 8 styles – the size of the stylistic gender gap between men and women on each topic. To do so, I

estimate models where I interact the gender of the MP delivering a speech with the topic proportions that pertain to that speech:

$$y_{i(j)}^s = \alpha + \beta^1 Woman_j + \sum_{k=2}^K \beta_k^2 \theta_{i,k} + \sum_{k=2}^K \beta_k^3 (Gender_i \cdot \theta_{i,k}) + \epsilon_{i(j)} \quad (\text{A.8})$$

I use the coefficients of this model to calculate estimated average differences between men and women on speeches devoted to each topic, which I denote as:

$$\delta_k^s = \begin{cases} \beta^1 & \text{if } k = 1 \\ \beta^1 + \beta_k^3 & \text{if } k \neq 1 \end{cases} \quad (\text{A.9})$$

The average difference in style  $s$  between men and women on speeches that are entirely devoted to topic one is given by  $\beta^1$  (i.e. the baseline), and  $\beta^1 + \beta_k^3$  captures the average gender difference in style on speeches entirely devoted to topic  $k$ . I denote the gender difference on each topic and style as  $\delta_k^s$ . This specification allows me to capture the aggregate differences between men and women's use of a style on each topic. Positive values for  $\delta_k^s$  indicate that women use the style more than men in a given topic, and negative values suggest that women use the style less than men in a given topic.

I then estimate a second set of regression models to capture, for each topic, the relationship between time and topic prevalence. To do so, I first multiply the number of words in each speech by the vector of topic proportions for that speech, giving me the weighted number of words dedicated to a given topic for each speech in the data. I then sum these topic-weighted word counts across all speeches within a given calendar month, and use the summed word counts as the dependent variable for regressions of the form:

$$y_t^k = \alpha + \gamma_k YearMon_t + \epsilon_t \quad (\text{A.10})$$

Here,  $y_t^k$  is the number of words on topic  $k$  in time period  $t$ , and  $\gamma_k$  captures the linear relationship between time and topic prevalence for topic  $k$ . Positive values of  $\gamma_k$  imply that topic  $k$  becomes more prevalent in parliamentary debate throughout the study period, and negative values suggest that the topic becomes less prevalent over time.

If the topical confounding argument is correct, then for a style like human narrative – where I observe average convergence between men and women over time – it must be the case that there is a negative relationship between the gender gap on that topic and the relationship between topic

and time. That is, topics where women use human narrative more than men (positive coefficient from equation A.8) should be becoming less prevalent over time (negative coefficient from equation A.10).

The topical-confounding hypothesis implies different relationships between topical gender-gaps and changes in topic prevalence over time for different styles. For instance, for human narrative, the main analysis shows that women are more likely to use this style in the early period of the data and less in the later period. For this style, topical confounding would occur if topics where women use narrative *more* on average than men (positive  $\delta_k^s$  from equation A.9) became *less* prevalent over time (negative  $\gamma_k$  from equation A.10), or the topics where women use narrative *less* than men (negative  $\delta_k^s$  from equation A.9) became *more* prevalent over time (positive  $\gamma_k$  from equation A.10). For human narrative, then, the topical-confounding hypothesis implies a negative relationship between the two sets of coefficients.

On the other hand, the aggregate results suggest that women are less aggressive than men in the early period of the data but are equally as aggressive later in the period. Accordingly, if this convergence can be explained by changes to the topics under discussion, it must be the case that the topics on which women tend to be less aggressive than men (negative  $\delta_k^s$  from equation A.9) become less prevalent over time (negative  $\gamma_k$  from equation A.10), or that the topics on which women tend to be more aggressive than men (positive  $\delta_k^s$  from equation A.9) become more prevalent over time (positive  $\gamma_k$  from equation A.10). Therefore, for aggression, the topical confounding hypothesis implies a positive relationship between the two sets of coefficients.

Following this logic through all eight styles, the topical-confounding explanation suggests that I should observe a positive relationship between  $\gamma_k$  and  $\delta_k^s$  for aggression, complexity, fact and negative emotion, and a negative relationship between  $\gamma_k$  and  $\delta_k^s$  for human narrative, affect, positive emotion, and repetition.

In figure A.8 I evaluate these expectations by plotting the estimated values of  $\gamma_k$  and  $\delta_k^s$  against each other for each style. In this plot, each point represents a single topic from the  $K = 40$  topic model: the x-axis measures the gender gap in the use of a given style ( $\delta_k^s$ ), and the y-axis measures the changing prevalence of the topic over time ( $\gamma_k$ ). I also fit a regression line between the sets of coefficients, which is coloured in red if the slope of the line is associated with a  $p$ -value of less than 0.05, and otherwise is coloured in grey.

The main implication of this analysis is straightforward: I find very little evidence to support the topical-confounding hypothesis. The size of the gender gap measured for a given style on a given topic largely does not predict the degree to which that topic becomes more or less prevalent over



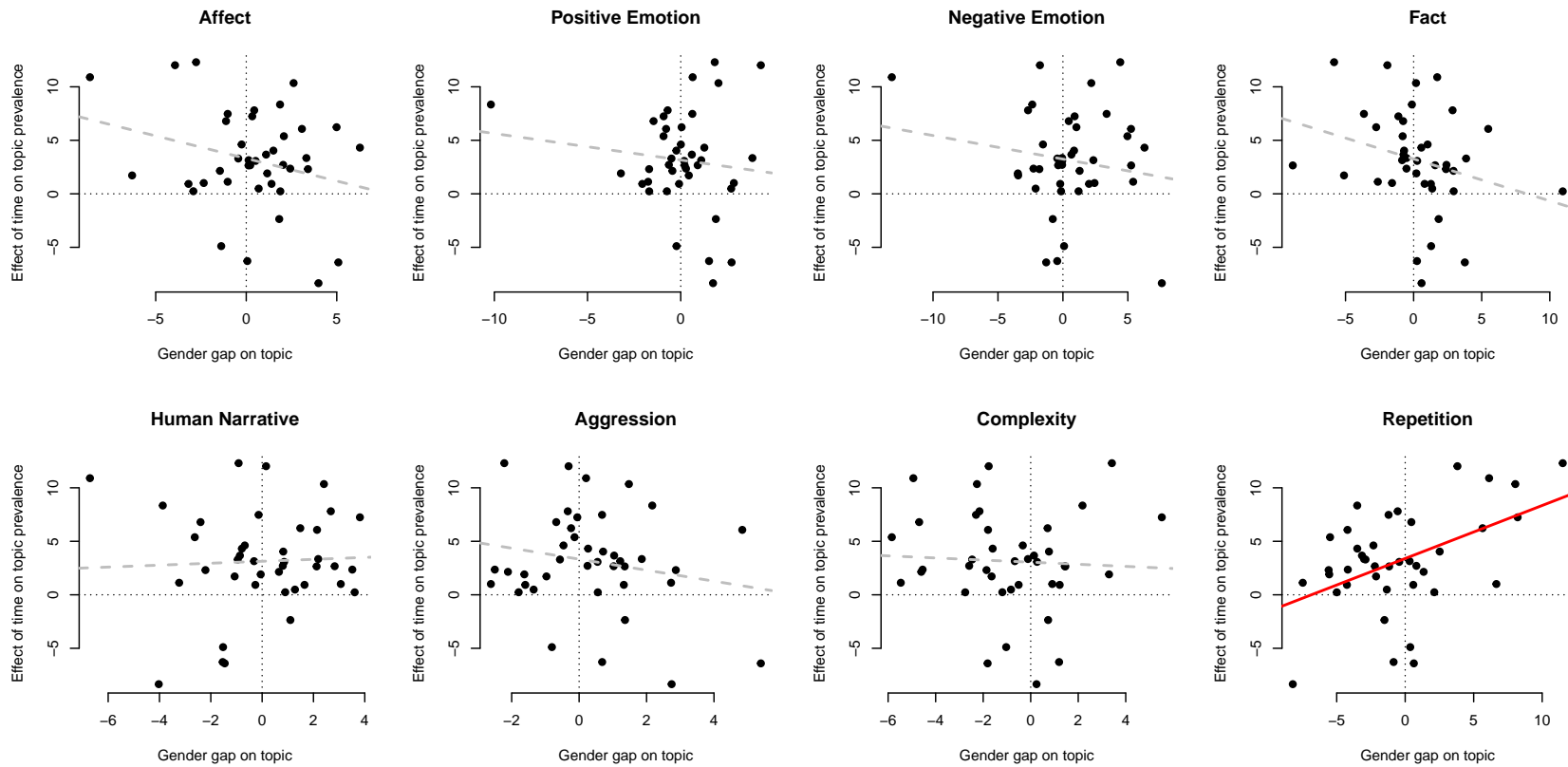


Figure A.8: **Topical-confounding**: The figure shows the relationship between the gender gap in the use of a given style on a given topic (x-axis), and the change in the prevalence of a given topic over time (y-axis).

time. For three of the styles – aggression, negative emotion, and fact – the relationships in figure A.8 are negative, where they would need to be positive for topical changes to explain the stylistic convergence I document in figure 3.2 of chapter 3. I also find a relationship that is in the “wrong” direction for repetition (that is, although statistically significant, the relationship would need to be negative to cause concern), and there is also essentially no relationship between the gender gap in human narrative on different topics and the changing prevalence of those topics over time. For the remaining styles – affect, positive emotion, and complexity – I do find some evidence that topics on which women display more of these styles become more prevalent over time, but the relationships are very noisy and in none of those cases am I able to reject the null hypothesis of a relationship of zero.

As there is no *a priori* reason to base my inferences on the  $K = 40$  topic model, in figure A.9 I summarise the relevant results from all 8 topic model specifications. In this plot, the x-axis measures the value for  $K$ , and the y-axis measures the slope of the regression line for the changing prevalence of a topic over time ( $\gamma_k$ ) as a function of the gender gap in the use of a given style in that topic ( $\delta_k^s$ ). The results clearly demonstrate that the findings are not sensitive to the number of topics used in the analysis. For all models, I find patterns that are very similar to those depicted in figure A.8. The only exception is that I find a significant coefficient for the “fact” style in the  $K = 80$  topic model. However, again, this relationship is in the “wrong” direction as it suggests a negative relationship between the topic-specific gender-gap in factual language and over-time topic prevalence, where the topical confounding story implies a positive relationship between these quantities for the factual language style.

Taken together, these analyses imply that the aggregate patterns I observe in chapter 3 cannot be convincingly explained by changes to the parliamentary agenda over time.

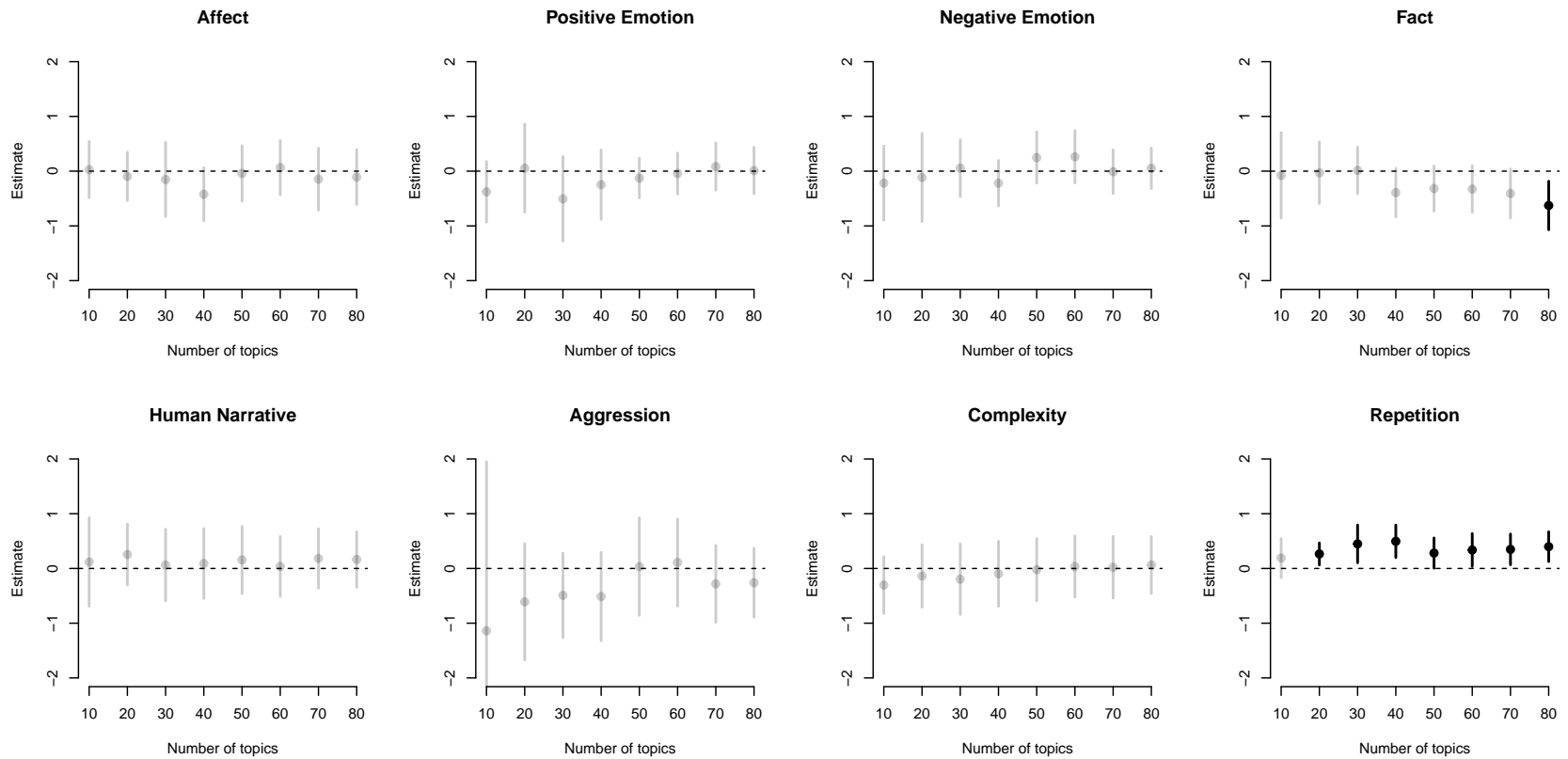


Figure A.9: **Topical-confounding, varying  $K$** : On the y-axis, the figure summarises the linear relationship between the gender gap in the use of a given style on a given topic ( $\delta_k^s$ ), and the change in the prevalence of a given topic over time ( $\gamma_k$ ). The x-axis measures the number of CTM topics,  $K$ , used to estimate these relationships.

## A.7 Style use and debate participation

My results show that, on average, women MPs deliver speeches that are less likely to be marked by communal styles and more by agentic style over time. One potential alternative explanation for these results is that men and women MPs who employ different speaking styles might have become differentially likely to *participate* in parliamentary debate over time. One might imagine, for instance, that women MPs who tend to deliver highly agentic speeches gave more speeches in parliament over the course of the study period, and that women who tend to deliver highly communal speeches participated less in debate over time. If that were the case, differential participation might drive the changing gendered speechmaking dynamics that I document, rather than within-MP changes.

To investigate this alternative explanation, I assess whether the average style of an MP across all speeches in a given parliamentary term predicts the number of speeches that the MP delivers. I begin by measuring the number of speeches delivered by each MP in each parliamentary term ( $\# \text{Speeches}_{i(t)}$ ), which I then model as a function of the gender of the MP, the average style of speeches given by the MP in that term ( $\text{Style}_{i(t)}^s$ ), and the interaction between these two variables. Specifically, for each parliamentary term,  $t$ , and each style,  $s$ , I estimate a model of the following form:

$$\# \text{Speeches}_{i(t)} = \alpha + \beta_1 \text{Woman}_i + \beta_2 \text{Style}_{i(t)}^s + \beta_3 (\text{Woman}_i \cdot \text{Style}_{i(t)}^s) + \epsilon_{i(t)} \quad (\text{A.11})$$

The key quantities of interest here are  $\beta_2$ , which measures the effect of a standard deviation increase in the use of a given style on the number of speeches delivered by men, and  $\beta_2 + \beta_3$ , which gives the same quantity for women MPs. If the results are driven by a selection-based story about the types of MPs who choose to participate in debate, then I should find that these two quantities broadly mirror the aggregate patterns I document in figure 3.3 of chapter 3. For example, if differential participation is the explanation for the decreasing average use of human narrative by women MPs, then I should observe a weaker relationship between the degree to which a woman MP's speeches tend to feature human narrative and the number of speeches delivered by that MP over time. Similarly, for negative emotion, if selection into debate drives the increasing use of that style by women, I would expect to see the relationship between the use of negative emotion and the number of speeches delivered by women MPs to have strengthened over time. I present the quantities of interest for each style in each parliamentary term in figure A.10.

In general, I find very little evidence that the average style of an MP predicts participation in debate

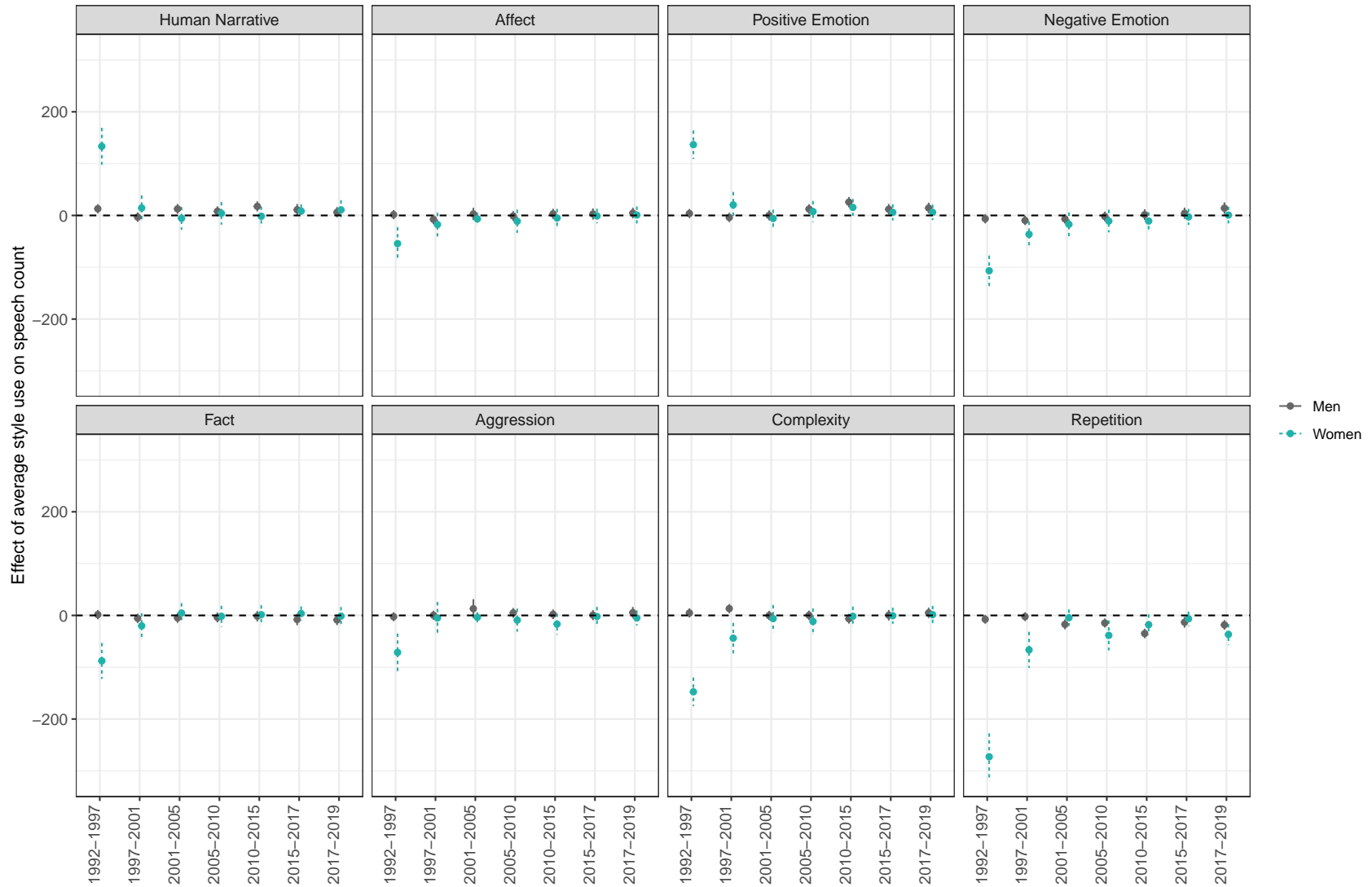


Figure A.10: **Participation as a function of average style use, by parliamentary term:** The figure illustrates the average marginal effect of a one standard deviation increase in the average style use on the number of times an MP speaks in a given parliamentary term.

at any point during the study period. Across almost all styles, the effects are indistinguishable from zero, implying that it is very unlikely that the results are driven by which MPs choose to speak in debate. Moreover, there are no clear over-time trends in these coefficients, which undermines the idea that, for example, women with more agentic speaking styles participate more over time. In other words, this analysis suggests that the sample of speeches that I observe do *not* appear to be disproportionately delivered by the more communal women MPs in the early period, and by more agentic women MPs in the later period. Rather, this analysis suggests that the changes over time that I document in chapter 3 are largely driven by within-MP changes in speaking style, and the replacement of MPs with different style-types over time (see figure A.7 above).

# Appendix B

## Appendix B – Earning Their Stripes? Gender, Political Experience, and Policy Prioritisation in Parliamentary Debate

### B.1 Seed words

In table B.1, I present the seed words for each of the policy areas which I measure in chapter 4.

Table B.1: Seed words

Issue	Seed words
Agriculture	farmland; farm; rural; farmers; agriculture; food; livestock; poultry; crop; dairy; seafood; pesticide; meat; grain; agricultural; cattle; organic; farms; farmer; beef; wheat; animal; animals; cotton; crops; fair-trade; free-range; breeding; fishing; environmental; environment; horticulture; horticultural; production; produce; land; tractor; breed; fertiliser
Children and Family	children; family; parent; parents; parenthood; families; mother; father; maternity; paternity; baby; care; couples; child; pregnancy; pregnant; childcare; mothers; fathers; childhood; youth; teenage; teenager; divorce; parenting; grandparent; parents; domestic; divorced; vulnerable; generation; single-parent; ageing; generations
Crime and Policing	crime; law; criminal; justice; fraud; corruption; illegal; trafficking; court; police; policing; prison; enforcement; fines; juvenile; violence; prisoner; murder; offences; punishment; riots; disturbances; asylum; courts; authorities; probation; prison service; counterterrorism; commissioners; cybercrime; detective; commissioner; violent; prosecution; prosecuted; homicides
Defence	defence; armed; forces; military; security; nato; intelligence; foreign; mobilisation; fleet; wartime; nuclear; disarmament; weapons; arms; manpower; army; navy; veteran; veterans; supplies; overseas; conflicts; war; terrorism; terrorist; alliance; terror; espionage; demobilisation; alliances; deterrent; trident; missile; treaty; landmine; atomic; proliferation; pow; submarine; civilian; civilians; personnel; maritime; bombs; bullets; homeland
Education	education; tuition; student; university; universities; elementary; secondary; grammar; school; schools; primary; teachers; pupils; class; scholarship; literacy; numeracy; examinations; knowledge; college; degree; educate; curriculum; development; pupil; children; school-aged; primary-age; secondary-age; classroom; higher education; enrolment; grades; assessment; educational; comprehensive; ofsted; schooling; academic; polytechnic; pedagogical; vocational; apprenticeship; postgraduate; graduate; gcse; gcses; a-levels

Issue	Seed words
Environment	pollution; environment; conservation; environmental; waste; hazardous; regulation; pollution; warming; climate; carbon; greenhouse; emissions; recycling; wildlife; coastal; protection; net-zero; sustainable; planet; sustainability; drought; defra; habitat; habitats; biodiversity; flood; floods; natural; resources; sustainable; re-use; recycle; development; climate change; global warming; greener; pollute; polluter; energy-efficient; food waste; net zero; flooding; natural disasters; biosafety; growth; deforestation; carbon dioxide; wildfires; decarbonisation
Finance and Economy	£; bank; budget; budgeting; deficit; growth; tax; taxes; treasury; economic; monetary; price; cost; inflation; interest; growth; monetary; macroeconomic; chancellor; jobs; wages; employee; employer; employers; interest; business; microeconomics; economy; income; recession; recovery; appropriations; credit; debt; finance; expenditure; fiscal; borrowing; taxation; loan; loans; banking; financial; regulation; commodities; investment; insurance; bankruptcy; deregulation; consumer
Health	nhs; health; hospital; doctor; medical; treatment; patients; drug; mental; healthcare; clinical; nursing; nurse; hospitals; doctors; nurses; hospitalised; rehabilitation; terminally; prescription; obesity; disease; gp; medicine; patient; counselling; psychotherapy; physiotherapy; psychiatry; disorders; cancer; diabetes; smoking; medication; medications; treatments; physicians; dentist; dental; vaccination; medic; paramedic; vulnerable; sickness; sick; ageing
Trade	trade; customs; export; import; regulations; tariff; negotiations; wto; domestic; international; markets; goods; services; products; treaty; treaties; brexit; europe; european; union; eu; borders; imports; exports; tariffs; trading; exporters; non-tariff; freight; cargo; shipping; agreements; free trade; trade deal; trade deals; commerce; infrastructure; consumers; customs union; cross-border; border controls; single market; customs check; deregulate; region; negotiators
Transport	transportation; transport; walking; cycling; bus; underground; highway; road; traffic; airports; airlines; aviation; railroad; rail; stations; passenger; automobile; driving; highways; drivers; trucks; lorries; maritime; routes; air; buses; roads; passengers; train; trains; station; commute; commuter; journey
Welfare	work; pension; services; housing; benefit; income; support; assistance; poverty; low-income; welfare; pensions; social; security; cuts; poverty-fighting; disabilities; disabled; poorer; poor; benefits; pensioner; jobcentres; claimants; universal credit; jobseekers; employment; unemployment; unemployed; payments; low-paid; working; living wage; minimum wage; zero-hour; allowance; wage; wages; austerity; tax; fraud; homeless; homelessness; food bank; food banks; debt; worklessness; workless; unemployment benefits; out-of-work; social security; social services; welfare state



## B.2 Validation tests

To validate the various measurement approaches used, I here provide a wide range of validation checks of which I describe each in turn.

### *Word-level validation*

In table B.2, I include the high-weight words (*Top*) for each of the issue areas, the words that are not included in the original seed dictionaries but are being given high-weights via the word-embedding approach (*Added*), and the words which are included in the original seed dictionaries but have nevertheless been given low-weights by the word-embedding approach (*Removed*). For all issue areas, the results are encouraging and show that words such as “patients”, “care”, “doctors” and “hospital” are the most highly-weighted for the health policy area, and words such as “farming”, “livestock”, “organic”, and “crops” for agriculture. Similarly, the approach seems to be successfully finding words that are not in the original seed dictionaries but are nonetheless deemed representative of the areas. For instance, words such as “troops”, “deployed”, “soldiers” and “peacekeeping” appear for defence, and words such as “qualification”, “subjects”, “classes”, and “teach” for the education policy area.

Agriculture			Children and Family			Crime and Policing		
Top	Added	Removed	Top	Added	Removed	Top	Added	Removed
farming	farming	breed	children	young	ageing	criminal	cases	detective
dairy	sheep	farmland	parents	adults	teenager	crime	crimes	disturbances
livestock	producers	grain	child	women	youth	courts	officers	riots
farmers	pig	environmental	mothers	carers	paternity	police	dealing	prisoner
production	fur	wheat	young	older	divorce	offences	offence	asylum
sheep	products	cotton	mother	daughter	divorced	cases	victims	juvenile
farms	fish	horticulture	families	lone	domestic	crimes	offenders	punishment
producers	processing	land	parent	disabled	childhood	justice	criminals	commissioner
pig	exports	horticultural	family	elderly	parenting	officers	serious	illegal
farm	steel	crop	adults	babies	childcare	prosecution	convicted	fines
beef	industry	breeding	women	wife	generation	enforcement	prosecutions	authorities
organic	milk	rural	fathers	grandparents	generations	dealing	alleged	corruption
food	imported	produce	carers	girls	maternity	court	involving	commissioners
fur	stocks	farmer	older	married	pregnancy	offence	organised	probation
agricultural	salmon	animal	father	son	baby	victims	immigration	trafficking
poultry	upland	fishing	couples	husband	pregnant	offenders	relation	murder
products	imports	cattle	daughter	caring	vulnerable	violence	prevent	prosecuted
fish	manufacturers	animals	lone	000-year-old	teenage	law	abuse	policing
crops	feed	agriculture	disabled	younger	care	criminals	investigations	violent
meat	abattoirs	meat	elderly	born	couples	serious	perpetrators	fraud
agriculture	fisheries	crops	babies	separated	father	convicted	terrorism	prison
animals	subsidies	poultry	wife	woman	fathers	prison	investigating	law
cattle	arable	agricultural	care	relatives	family	fraud	investigation	violence
processing	plant	food	teenage	men	parent	violent	cps	court
fishing	coal	organic	grandparents	marriage	families	policing	acts	enforcement
animal	industries	beef	girls	teenagers	mother	prosecutions	act	prosecution
exports	gas	farm	married	children's	mothers	alleged	secondly	justice
farmer	tourism	farms	son	adult	child	prosecuted	sexual	offences
steel	plants	production	husband	victims	parents	involving	victim	police
industry	puppy	farmers	caring	home	children	organised	prosecute	courts

Defence			Education			Environment		
Top	Added	Removed	Top	Added	Removed	Top	Added	Removed
military	afghanistan	bullets	pupils	teaching	elementary	environment	marine	polluter
forces	iraq	wartime	education	qualifications	enrolment	environmental	energy	greener
nuclear	capability	alliances	schools	students	educate	biodiversity	water	energy-efficient
armed	operations	homeland	teaching	training	knowledge	climate	food	decarbonisation
weapons	troops	veteran	school	colleges	development	marine	impacts	drought
personnel	capabilities	atomic	qualifications	maths	tuition	energy	gas	recycle
defence	deployed	bombs	teachers	teach	schooling	pollution	reduce	defra
afghanistan	iraqi	disarmament	curriculum	nursery	examinations	sustainable	global	hazardous
iraq	threat	treaty	secondary	courses	assessment	emissions	reducing	planet
army	deployment	alliance	students	taught	numeracy	conservation	targets	floods
capability	air	fleet	academic	qualification	class	waste	impact	warming
nato	force	conflicts	training	a-level	grades	development	efficiency	regulation
operations	conventional	manpower	primary	higher	graduate	water	management	habitat
troops	chemical	terror	vocational	apprenticeships	degree	carbon	safety	growth
capabilities	british	supplies	colleges	post-000	comprehensive	wildlife	agency	coastal
security	united	trident	university	skills	ofsted	sustainability	tackling	habitats
missile	naval	submarine	maths	learning	apprenticeship	flood	renewable	resources
deployed	equipment	arms	teach	academies	gsces	food	areas	recycling
iraqi	russian	deterrent	educational	parents	grammar	impacts	countryside	flooding
threat	soldiers	overseas	universities	000-year-olds	college	gas	strategy	greenhouse
deployment	iran	foreign	nursery	girls	a-levels	reduce	rural	protection
air	international	maritime	courses	subjects	literacy	natural	developing	natural
force	combat	proliferation	taught	attainment	postgraduate	global	protect	flood
navy	uk	veterans	children	places	student	protection	thirdly	sustainability
conventional	peacekeeping	intelligence	qualification	level	pupil	greenhouse	infrastructure	wildlife
chemical	attack	terrorism	a-level	mathematics	classroom	reducing	stocks	carbon
civilian	syria	terrorist	gcse	science	gcse	flooding	measures	development
war	conflict	civilians	higher	study	children	targets	uk's	waste
british	weapon	war	apprenticeships	classes	universities	impact	protecting	conservation
united	deploy	civilian	post-000	youngsters	educational	recycling	resource	emissions

Finance and Economy			Health			Trade		
Top	Added	Removed	Top	Added	Removed	Top	Added	Removed
investment	secondly	commodities	patients	care	physicians	eu	uk	cargo
tax	costs	bankruptcy	treatment	gps	ageing	trade	market	deregulate
financial	moreover	budgeting	care	professionals	sickness	uk	uk's	commerce
secondly	billion	deregulation	mental	psychiatric	vaccination	union	united	negotiators
costs	sector	employee	nhs	acute	physiotherapy	european	countries	region
cost	increase	employer	health	services	disorders	europe	britain	freight
borrowing	thirdly	recession	cancer	specialist	vulnerable	markets	industry	shipping
moreover	capital	monetary	gps	specialists	terminally	market	thirdly	tariff
£	rate	wages	doctors	stroke	sick	uk's	deals	infrastructure
growth	furthermore	loan	medical	practitioners	obesity	united	competition	cross-border
billion	rates	recovery	dental	clinicians	smoking	export	secondly	exporters
sector	increases	banking	patient	illness	dentist	agreements	british	treaties
increase	pay	regulation	professionals	consultants	counselling	exports	furthermore	brexit
debt	businesses	employers	psychiatric	beds	prescription	countries	companies	borders
thirdly	benefit	treasury	acute	in-patient	medication	britain	ec	wto
income	whereas	consumer	nurses	trusts	disease	industry	kingdom	domestic
taxation	lower	fiscal	hospital	dentists	drug	international	global	services
economy	cut	price	nurse	staff	diabetes	thirdly	states	regulations
taxes	overall	loans	treatments	dementia	medicine	trading	regulation	products
capital	year	bank	services	diagnosis	rehabilitation	goods	operate	customs
rate	higher	jobs	gp	children's	nursing	deals	deal	consumers
finance	savings	deficit	hospitals	surgery	healthcare	competition	existing	treaty
furthermore	companies	expenditure	specialist	condition	clinical	secondly	negotiate	import
economic	meant	interest	doctor	breast	doctor	british	eu's	imports
budget	mean	inflation	specialists	elderly	hospitals	negotiations	britain's	tariffs
rates	current	insurance	stroke	treating	gp	furthermore	moreover	negotiations
increases	spending	credit	clinical	dentistry	treatments	companies	developing	goods
business	means	business	practitioners	carers	nurse	ec	agreement	trading
pay	terms	budget	clinicians	service	hospital	kingdom	means	international
credit	level	economic	healthcare	midwives	nurses	tariffs	including	exports

<b>Transport</b>			<b>Welfare</b>		
Top	Added	Removed	Top	Added	Removed
rail	freight	commute	benefit	families	workless
bus	railway	trucks	benefits	pensioners	worklessness
buses	network	highway	families	help	jobseekers
freight	railways	driving	employment	credit	jobcentres
road	operators	highways	income	people	out-of-work
transport	journeys	aviation	pensioners	households	austerity
passengers	airport	maritime	help	credits	fraud
passenger	travelling	airlines	credit	living	homeless
traffic	travel	journey	disabled	child	assistance
trains	london	transportation	housing	claiming	disabilities
train	motorway	cycling	people	pay	debt
roads	vehicles	walking	social	helping	homelessness
railway	cars	lorries	claimants	incomes	security
routes	infrastructure	stations	households	disability	poorer
network	high-speed	drivers	working	increase	low-paid
railways	ferry	station	tax	secondly	wages
operators	heathrow	air	poverty	moreover	low-income
commuter	tube	underground	credits	receive	services
journeys	ports	airports	living	furthermore	pensioner
airports	route	commuter	child	children	pension
underground	services	routes	work	receiving	poor
airport	connectivity	roads	allowance	addition	unemployed
air	fares	train	claiming	costs	unemployment
travelling	users	trains	pensions	reduced	cuts
travel	commuters	traffic	pay	paid	support
london	car	passenger	helping	cut	wage
motorway	congestion	passengers	incomes	helps	payments
station	m000	transport	disability	older	welfare
vehicles	capacity	road	increase	means	pensions
drivers	a000	buses	welfare	receipt	allowance

Table B.2: Word-level validation

*Sentence-level validation*

Next, in tables B.3–B.6 below, I show the 10 sentences with the highest scores for each of the issue areas for which I apply the word-embedding based approach. For all issue areas, the sentences seem to reflect the issues of interest. For instance, the sentences on agriculture refer to issues of farming and various food industries. Similarly, the sentences for the education issue area refer to issues of schooling, examination, further education, and qualifications. Transport is similarly represented by appropriate topics such as public transport, cycling, or improving transport links. Overall, the top sentences provide reassurance that the word-embedding based approach is picking up on the appropriate topics for each of the issue areas.

Table B.3: Top sentences for Agriculture, Children and Family, and Crime and Policing

Agriculture	Children and Family	Crime and Policing
The same applies to arable farmers, beef producers, the pig industry and poultry farmers.	Among pensioners, single women lose 12.5% compared with single male pensioners losing 9.5% and pensioner couples losing 8.6%.	They include antisocial behaviour orders, child safety orders, local child curfews and statutory crime prevention partnerships between local authorities and the police.
Exceptions Food, agricultural and horticultural produce, fish and fish products, seeds, animal feeding stuffs, fertilisers and pesticides.	A baby boy born in Newham is likely to die nearly six years earlier than a baby boy born in Westminster.	The Crime and Disorder Act 1998 introduced measures to deal with antisocial behaviour, including the crime and disorder partnerships and antisocial behaviour orders.
Alongside tourism, and the production of planes, cars and steel, my area produces sheep and beef products that are exported.	They provide support to fathers, mothers, grandparents, foster parents and adoptive parents, and many different services.	In 2009, offenders subject to community orders committed more than 18,000 serious violent and sexual offences, including 172 sexual offences against children.
Look at the poultry industry, the pig industry, the horticulture industry or fruit and veg producers.	Now, someone in a low-paid job will receive child benefit, child credit, child care tax credit and working families tax credit.	We are also establishing a new drug prevention advisory service to support drug action teams on drug prevention, including prevention of drug-related crime.
Regulations apply to the production of food, animal feed and livestock, including bees and farmed fish, marketed as organic.	Many cancer patients receive treatment for more than a year, and face losing their employment and support allowance while still receiving treatment.	The CPS is prosecuting and convicting more defendants of domestic abuse, rape, sexual offences and child sexual abuse than ever before.
The products of our steel industry supply the booming automotive manufacturing industry and the aerospace manufacturing industry, among others.	Social services departments could furnish information about residential and nursing homes, people receiving home help and attendance allowance, and vulnerable young adults.	What recent steps the Crown Prosecution Service has taken to ensure that prosecutors are able more effectively to prosecute cases of domestic abuse.
Like many farmers across the country, dairy farmers do not trade directly with supermarkets but deal with processors and food manufacturers.	The two biggest groups that are affected are lone parents, particularly lone mothers, and single women pensioners.	The Crown Prosecution Service continues to improve its response to cases involving rape allegations and other forms of serious sexual offending.
Denmark is probably more focused on organic crops, but the UK has many producers in a niche market of organic foods.	They include adoption rights for same-sex couples, housing rights for same sex couples, pension rights for same sex couples.	What steps the Crown Prosecution Service has taken to enable its prosecutors effectively to prosecute stalking and harassment cases.
For the milk sector, lower intervention prices are likely to feed through to lower farm gate milk prices.	The prevalence of mental health issues among children living outside married family relationships is 75% higher than among children of married parents.	The Crown Prosecution Service and the Serious Fraud Office regularly engage with Scotland's prosecution service and the Public Prosecution Service for Northern Ireland.
The marketing and promotion of British meat and dairy produce is essential for our domestic and export markets.	It is true that a baby boy born in Manchester today will live seven years less than a baby boy born in Dorset.	The Crime and Disorder Act 1998 introduced crime and disorder partnerships, antisocial behaviour orders and changes to the youth justice system.

Table B.4: Top sentences for Defence, Education, and Environment

Defence	Education	Environment
This capability includes delivering maritime force protection, providing maritime security to deployed forces, global maritime reach and maritime support to civil organisations.	Core schools would provide an alternative key stage 4 curriculum, with English, maths and science alongside two further technical qualifications.	Energy from waste cannot contribute towards waste reduction, but does contribute towards more sustainable waste management and the UK's landfill directive targets.
Control of Weapons Control of nuclear, biological and chemical weapons and other weapons of mass destruction.	Why should new institutions like free schools, academies and UTCs be able to provide free school meals while FE and sixth-form colleges cannot?	Our key objectives include the sustainable use of natural resources and protecting the rural, urban, marine and global environments.
Future carriers will be capable of deploying air power in the form of the joint Royal Navy-Royal Air Force future carrier-borne aircraft.	Four out of five of our young children are getting their early-years education experience in nursery schools, nursery classes or reception classes.	It could create natural barriers that would enhance soil retention and improve river water quality, while increasing the area of wildlife habitats.
Iran already has missiles capable of carrying nuclear warheads, and our Prime Minister warned in 2012 that Iran is developing intercontinental ballistic missiles.	Free schools are providing parents with choice, not just in mainstream education but in ensuring high-quality provision for children with special educational needs.	Sustainable development is not only a question of biodiversity, environmental protection, traffic reduction, renewable energy and securing clean air and water.
More than 500 UK personnel are on UN-led peacekeeping missions, in addition to the 7,000 personnel on UN-mandated missions in Bosnia and Kosovo.	More students took A-levels in FE colleges than in schools and more students went to university from FE colleges than from schools.	It may help to mitigate the effects of climate change by promoting the development of offshore renewable energy and carbon capture and storage.
Emphasis is laid on regional wars, aid and peacekeeping, nuclear proliferation, terrorism, computer security and chemical and biological weapons.	On a more basic level, last year colleges helped about 0.3 million adults to improve their basic literacy and numeracy skills.	The plans will set targets and identify actions to develop public transport, improve air quality, enhance road safety and reduce road traffic.
The maritime presence currently in theatre consists of 30 Royal Navy and Royal Fleet Auxiliary ships and two Tomahawk cruise missile-capable submarines.	In January 2001 there were 65,500 teaching assistants - that includes classroom assistants - employed in maintained nursery and primary schools in England.	They can help to reduce greenhouse gas emissions, improve local air quality and reduce noise pollution.
As of today, 900 Royal Navy personnel and 1,800 Royal Air Force personnel are also deployed in support of operations in the Balkans.	It has a college of further education, a sixth-form college, a specialist art and design college and a Catholic school sixth-form college.	That covers sustainable farming-reducing pollution and greenhouse gas emissions, adapting to climate change and protecting against pests and diseases-and tackling food waste.
Iran, with foreign assistance, is buying and developing longer-range missiles, already has chemical weapons and is seeking nuclear and biological capabilities.	They are getting better results - 56.7 per cent. of pupils in specialist schools get five good grades at GCSE.	Current resources are not sufficient to meet fuel poverty commitments or wider energy efficiency targets set out in the clean growth strategy.
We demand the destruction of Iraq's capacity to deliver weapons of mass destruction of any kind - nuclear, biological or chemical.	We have provided much better non-repayable grants and support for students with disabilities in post-16 and higher education, including part-time and postgraduate students.	That could include improving habitats, water quality and soil health, promoting biodiversity, advancing animal welfare and allowing public access.



Table B.5: Top sentences for Finance and Economy, Health, and Trade

Finance and Economy	Health	Trade
Public sector pay increases generate tax revenues, reduce social security expenditure, inject extra value into the economy and create jobs.	We have specialised children's palliative care commissioning, general children's palliative care commissioning and the commissioning of children's social services.	Import tariffs for agricultural goods entering the UK from outside the EU are EU wide and subject to World Trade Organisation rules.
It must mean lower business taxes, lower business rates and better access to credit, and it must also mean substantially less regulation.	That means that our cardiac unit, our stroke unit, our intensive care unit and our accident and emergency unit will close.	Failing that, even single market access via the European economic area and customs union membership would also allow current arrangements to continue unhindered.
Tax increment financing schemes allow local authorities to fund major regeneration projects by borrowing against the future increase in local business rate income.	They are the Colchester Hospital University NHS Foundation Trust, the Dudley Group NHS Foundation Trust and the Blackpool Teaching Hospitals NHS Foundation Trust.	United Kingdom airlines operating out of United Kingdom airports - Newcastle airport provides a very good example - provide services for foreign airlines.
Rising energy bills and fuel bills are coming on top of higher taxes, cuts to tax credits and cuts to public services.	There will be better emergency care, shorter waiting times and improvements in cancer, heart, mental health, children's and elderly services.	The Government's UK Trade and Investment organisation helps UK companies to win export business globally and overseas businesses seeking to invest in the UK.
Helping people in work to earn higher wages means lower benefit payments and, again, greater income tax and national insurance receipts.	Patients may require referral to specialist mental health services for diagnostic assessment or advice about medication or other aspects of their care.	Prices are rising mainly as a result of increasing oil prices, leading to higher wholesale gas prices, and higher international coal prices.
Failing to act now would risk higher interest rates, higher mortgage rates, higher rates of business failure and higher unemployment.	It refers to primary care groups or primary care trusts commissioning directly from a private or voluntary health care provider.	It includes liberalisation in trade in services well beyond World Trade Organisation commitments and building on recent EU free trade agreements.
The small firms loan guarantee scheme and the enterprise finance guarantee scheme have produced less than half the target amount of lending.	Evidence-based care and treatment are provided across community-based settings through support from primary care, with specialist or hospital in-patient services provided as appropriate.	It will also improve the regulation of consumer credit businesses and establish a fair and competitive framework for consumer credit agreements.
Stronger growth and sound public finances are the only sustainable way to deliver better public services, higher real wages and increased living standards.	They are all running diabetic clinics, chronic obstructive pulmonary disease clinics, asthma clinics and minor surgery clinics.	Multilateral environmental treaties, like the climate change convention, should no longer be open to endless challenge under World Trade Organisation rules.
What we do for business is maintain conditions of economic stability, low interest rates and low inflation while introducing measures to stimulate investment.	NHS trusts and primary care trusts will have patient advice and liaison services, providing on-the-spot help and information about health services.	Scotland's vibrant financial services sector is important to the UK's competitiveness, and more particularly to Scotland's competitiveness in the global marketplace.
A fund manager's ability to pay capital gains instead of income tax allows them to avoid paying national insurance on part of their income.	Older people have very different physical, social and psychological issues, which require specialist old-age psychiatrists working in specialist services for older adults.	It means leaving the customs union and the single market, and leaving the common agricultural policy and the common fisheries policy.

201

Table B.6: Top sentences for Transport and Welfare

Transportation	Welfare
The sources extend right across the spectrum of vehicles, including buses, taxis, heavy goods vehicles, light goods vehicles and cars.	Crime doubled, national debt doubled, those on long-term income support doubled, long-term unemployment doubled, homelessness doubled and child poverty doubled.
The crowded west coast main line currently combines long-distance inter-city, inter-regional and commuter passenger services, together with freight.	It also increases spending on disabled households and enables disabled people to earn nearly £8,000 a year without affecting their universal credit entitlement.
It would mean significantly more commuter services, better connectivity and more routes for rail freight, taking lorries off our most congested roads.	Now, someone in a low-paid job will receive child benefit, child credit, child care tax credit and working families tax credit.
We need to link bus routes with the London underground network to reduce car use in south-east London.	The minimum income guarantee linked to wages means that there is a growing gap between the basic state pension and the guaranteed minimum income.
We should promote freight services on the channel tunnel rail link via the north London connection to the west coast main line.	Rising energy bills and fuel bills are coming on top of higher taxes, cuts to tax credits and cuts to public services.
One project being discussed as part of the surface access plans is the southern rail access project to improve rail links to Heathrow airport.	Among pensioners, single women lose 12.5% compared with single male pensioners losing 9.5% and pensioner couples losing 8.6%.
Substantial congestion exists on the M60, M62 and M56 around Greater Manchester, but Highways England's investment plans start north-east of this area.	Including housing benefit and council tax benefit, the proportion of UK pensioners entitled to means-tested benefits last year was around 50 per cent.
Cyclists use cycle lanes, and motorists and other road users use the roads alongside them.	Cuts to the severe disability premium mean that disabled lone parents with young carers stand to lose £58 a week.
The new line would connect the planned HS2 and Crossrail interchange station at Old Oak Common with the West Coast Main Line."	We now spend nearly £11 billion extra each year on pensioners, and almost half that additional spending goes to the poorest third.
Reading is a major hub on the rail link connecting the main east-west routes from London to the main north-south routes from Birmingham.	Helping people in work to earn higher wages means lower benefit payments and, again, greater income tax and national insurance receipts.

### *Position and committee validation*

Next, if the measurement approach is identifying the relevant topics for each of the issue areas, then we would expect that politicians who hold a relevant ministerial position in either the government or opposition or sit on a relevant committee would talk more about a given issue area. To assess whether this is the case, I created a series of dummy variables that match each of the issue areas described above for whether any given MP holds a relevant ministerial role for either the government, opposition, or sits on a relevant committee at the point of speech delivery between 1997–2019.

To create these variables for government positions, I created a series of vectors collecting the names of various ministerial positions over time that are relevant to each issue area. The vectors of position names reflected changes in the names of roles and departments over the period of study. For instance, over the period of study, relevant titles for the education issue area have ranged from “Secretary of State for Education”, “Secretary of State for Children, Schools and Families” or “Secretary of State for Education and Science”. For each MP at any given point of time, they were marked as either holding a relevant government position for a policy area or not. To create these variables for opposition positions, I followed the same steps but identified all the relevant opposition positions across the time period. Finally, the same steps were taken for relevant committees. I identified all historic committee names that were relevant to the issue areas, and created a series of dummy variables for whether any MP was serving on a relevant committee at any point across the study period.

To assess whether MPs that hold relevant positions or sit on relevant committees talk more about a given issue area, I estimate a series of models for the topic-based outcomes  $Y_{id}^t$  for an individual MP  $i$  in a debate  $d$  for topic  $t$  of the following form:

$$Y_{id}^t = \alpha + \beta_1 GovMinister_i + \beta_2 OppMinister_i + \beta_3 CommMember_i + \epsilon_i \quad (B.1)$$

where  $\beta_1$  in each model describes the difference between MPs who are and are not government ministers relevant to an issue area,  $\beta_2$  describes the same quantity for opposition ministers, and  $\beta_3$  describes the same quantity for committee members. I cluster standard errors at the MP level  $i$ .

Figure B.1 shows the results from this analysis. Each panel of the figure represents a different issue area, as indicated by the panel titles, and shows the coefficients and 95% confidence intervals for opposition ministers, government ministers, and committee members. While the size of the effect varies depending on the issue and position,  $\beta_1$ – $\beta_3$  are always positive, indicating that compared to those who do not hold these positions, MPs who hold relevant positions do talk more about the issue

areas. Or, alternatively, an MP who holds a government ministerial position related to health, talks significantly more about health than an MP who does not hold such a position. Overall, therefore, the results from this validation analysis provide further support that the measurement approach is successfully detecting the issues that politicians talk about.

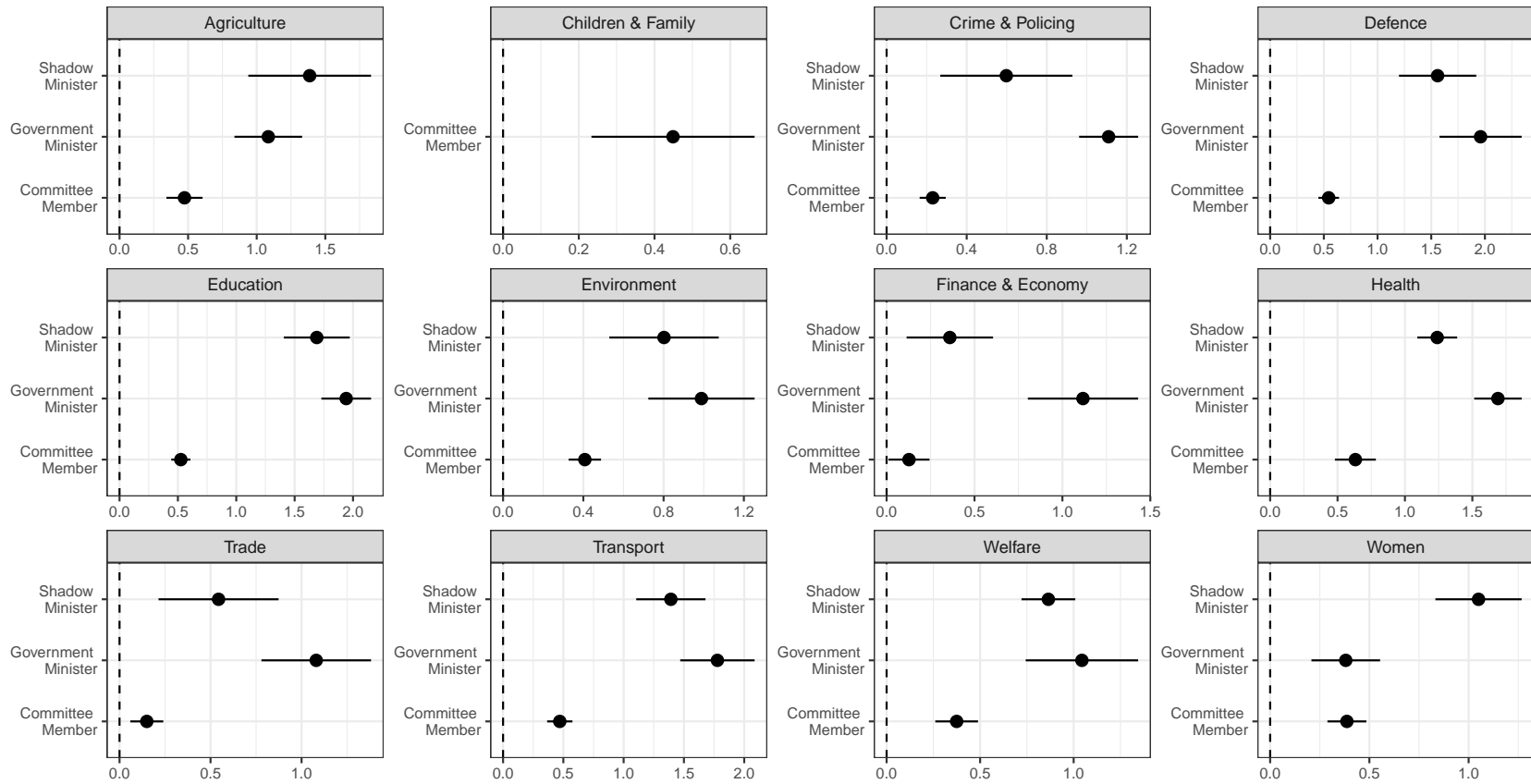


Figure B.1: Average topic use by government, opposition, and committee positions

*Debate-level validation*

The word-embedding based approach provides me with a measure of the degree to which each individual contribution by an MP in a given debate discusses each of the issue areas. A final source of reassurance that the approach is working as intended is to establish that speeches that pay particular attention to certain issues are also nested in debates on the same issues. For example, whether speeches that raise many words associated with welfare are delivered in debates that are also on welfare issues more broadly. I assess the debate level validity of these measures both qualitatively and quantitatively. Qualitatively, I aggregate the word-embedding scores from the individual in a debate level to the debate level to identify the degree to which a debate *overall* is about a given issue area. The structure of debate titles in Hansard means that at times answers to individual questions by MPs are classified as debates. To account for this, I drop question time debates and then extract the titles of the top 20 scoring debates for each of the issue areas. These are presented in tables [B.7–B.10](#).

In some cases, the titles are quite non-specific, such as simply “Petition” or “Committees”, which reflects the way that parliamentary debates are titled in Hansard. However, overall, these debate titles seem reassuring. For instance, the highest scoring finance and economy debates include titles referring to tax, incomes, benefits, and allowances. Similarly, the debates for children and family refer to adoption, fostering, young fathers, and sex education in schools. Therefore, qualitatively at least, debates that contain many speeches on a given issue area seem to be nested in well-matched debates topically.

Table B.7: Top debates for Agriculture, Children and Family, Crime and Policing

Agriculture	Children and Family	Crime and Policing
Petition: Factory Farming	Identification and Support of Carers	Domestic Violence
Petition: Clocaenog Wind Farm	Adoption and Fostering	Solicitor-General
Food Labelling	Point of Order	Restraining and Protection Orders
Animal Welfare	Petition: Carers	Petition: Sentencing
Farm Produce	Northern Ireland	Clause 22
Deferred Division	Young Fathers	Petition
Royal Assent	Older People's Rights Commissioner	Sexual Offences
Petition	Petition - Funding for Mental Health Provision in Cumbria	Alcohol-Related Anti-Social Behaviour and Domestic Violence
Private Members' Bills	Clause 70	Royal Assent
Poultry Industry	Sex Education in Schools	Victims
Sustainable Energy Agency	Draycott Manor School	Schedule 6
Petition	Petition: Education Maintenance Allowance	Royal Assent
Transport of Farm Animals	Care of Prisoners' Children	Bill Presented
Food Labelling	Petition	The Solicitor-General
Food Security Strategy	Adoptive Parents	Bill Presented
Petitions: Animal Welfare	Riverside Mental Health Care Unit	Petition: Minimum Sentence for Murder
Deferred Division	Parental Contact	Royal Assent
Movement of Farm Animals	Petition	Clause 22
Pet Crematoria	Orders of the Day: Lone Parents	Stalking
Live Animal Exports	Petition - the Future of Maintained Nursery Schools	Petition: Gaza

Table B.8: Top debates for Defence, Education, and Environment

Defence	Education	Environment
Royal Assent	Petition - the Future of Maintained Nursery Schools	Petition: Middleton on Sea
Nato	Petition: Funding for grammar schools in Southend	Doorstep Recycling Schemes
Petition - Protection for British Service Personnel	Draycott Manor School	Sustainable Energy Agency
Royal Assent	Schools and Colleges	Orders of the Day: Aluminium Recycling
Petition: Missile Defence	Petition - the Future of Nursery Schools	Petition
Palestine	Young Apprenticeships	Domestic Energy Efficiency
Anti-missile Defence	Petition: Closure of Downhills Primary School, Tottenham	Biofuels
Nuclear Global Threat	Free School Funding	Domestic Energy Efficiency Plan
Iraq	Orders of the Day: School Swimming	Bill Presented
UN Peacekeeping Week 2016:	Petition: Crowlands Junior School	Orders of the Day: Renewable Energy
International Assistance Force	Petitions: Further Education Funding	Peak Oil
Missile Defence	Catholic Sixth-form Colleges:	Petition
Petition - Uk Arms Export to Saudi Arabia	Colchester Sixth Form College	Renewable Energy
Royal Assent	Petition: Schools	Climate Change
Afghanistan	Petition - Funding for Mental Health Provision in Cumbria	Renewable Energy
Royal Assent	Telford Co-operative Multi Academy Trust Schools:	Orders of the Day: Flood Plains
Royal Navy Base: Bahrain	Petition: Education and School Equipment	Petitions: Consumers for Health Choice
International Terrorism	Schools	Future of Biomass
Afghanistan	Foreign Language Teaching	Energy
Iraq	Careers Guidance	Municipal Waste Recycling Bill



Table B.9: Top debates for Finance and Economy, Health, and Trade

Finance and Economy	Health	Trade
Community Finance Tax Credit	Petition: NHS Health Services in Guisborough, Skelton, Brotton and Park End	Orders of the Day: European Community Documents
Committees	Primary Care	Orders of the Day: European Community Documents
European Community Documents	Identification and Support of Carers	Orders of the Day: European Community Documents
Sustainable Energy Agency	Petition - Urgent Treatment Centre, Westmorland General Hospital	Industrial Development Bill
Orders of the Day: Ways and Means	Petition: NHS health services in Guisborough, Skelton, Brotton and Park End	Orders of the Day: European Community Documents
Pension Charges	Petitions: Herceptin	Orders of the Day: European Community Documents
Orders of the Day: Ways and Means	After Clause 64	Orders of the Day: European Community Documents
Universal Funded Pensions	Petition	Orders of the Day: Delegated Legislation
Petition - Tobacco Levy	Community Hospitals	Sustainable Energy Agency
Orders of the Day: Ways and Means	Asthma	Bill Presented
Introduction	Riverside Mental Health Care Unit	Petition: Transatlantic Trade and Investment Partnership
Introduction	Secondary Breast Cancer	Petition: Factory Farming
Orders of the Day: Ways and Means: Budget Statement	Cancer Care	Nato
Petition: Train Services	Petitions: Endometriosis SHE Trust	Orders of the Day: Delegated Legislation
Orders of the Day: Ways and Means	Petition: Dudley Borough Walk-in Centre	Orders of the Day: Delegated Legislation
Employment	Acute and Emergency Services	Poultry Industry
Small and Medium-Sized Enterprises	Palliative Care for the Terminally Ill	Orders of the Day: Delegated Legislation
Introduction	Point of Order	Orders of the Day: Delegated Legislation
Pensions	Bills Presented	Orders of the Day: Delegated Legislation
Planning	Petition - Funding for Mental Health Provision in Cumbria	Delegated Legislation

Table B.10: Top debates for Transportation and Welfare

Transportation	Welfare
Petition - Ruabon Station	Universal Funded Pensions
Train Services	Committees
Cotswold Railway	Petition - Funding for Mental Health Provision in Cumbria
Petition: Heaton Chapel Train Station	Petition
Petition - Speed Limit in Spatham Lane, Streat Lane and Underhill Lane, East Sussex	Domestic Energy Efficiency
Royal Assent	Income Tax
Petition: Traffic Noise	Asylum Seekers
Bromsgrove Station	Orders of the Day: Housing Costs
Petition: High Speed Rail	Pensions
Petition - Residential Development on Addington Road, Irthlingborough	Travel Concessions
Petition - M42 Motorway Service Area	Winter Fuel Payments
Eurorail Freight Route	Orders of the Day: Lone Parents
Petition: Train Services	Petition: Carers
A1	Orders of the Day: Ways and Means
Speed Limits	Inner Cities
Improvement of Rail Passenger Services	Orders of the Day: Ways and Means
Petition: Community Rail	Clause 5
Petition - Preston Park Train Services	Orders of the Day: Learning and Skills Bill
Sue Penney	Disabilities, Poverty and Inequalities:
Petitions: Rail Services	Housing

Quantitatively, I also assess whether MPs deliver more speeches on issues in debates that are titled with the same issue areas. To do so, I create a measure of each of the issues at the debate level using the debate titles. To identify whether the subject of a debate is focused on a given issue, I apply the seed words for each issue area to all debate titles and identify the number of words from each dictionary that appear in any given title. For each issue, I then transform these measures into a series of dummy variables where a debate is characterised as about a given issue if one or more of the seed words appear in the title.

With these measures in hand, I can assess, for instance, whether MPs deliver more speeches on issues of healthcare in debates that include healthcare specific titles. To assess whether this is the case, I use the dictionary created dummy variables of whether a debate is or is not about a given issue area, and estimate a series of models for the topic-based outcomes  $Y_{id}^t$  for an individual MP  $i$  in a debate  $d$  for topic  $t$  of the following form:

$$Y_{id}^t = \alpha + \beta_1 \text{DebateTopic}_d + \epsilon_i \quad (\text{B.2})$$

where  $\alpha$  in each model describes the average discussion of a given issue by MPs in a debate not characterised as about the issue, and  $\beta_1$  describes the difference in the discussion of a given issue by MPs in a debate that is about an issue compared to one that is not.

Figure B.2 presents the results from this analysis. Each panel again represents a different issue area, as indicated by the panel titles, and shows the relevant debate topic coefficient and 95% confidence intervals for. The panel for welfare, for instance, should be read as “how much more do MPs talk about welfare in speeches in a debate on the topic of welfare, compared to debates on other topics”. Reassuringly, across all issue areas, MPs talk significantly more about the issue if the debate is on the same issue. Further, the size of the effects for this analysis are large. Recall, the measures are standardised to have mean zero and standard deviation one. The size of the effects vary between approximately 0.25 and 2.5 standard deviations.

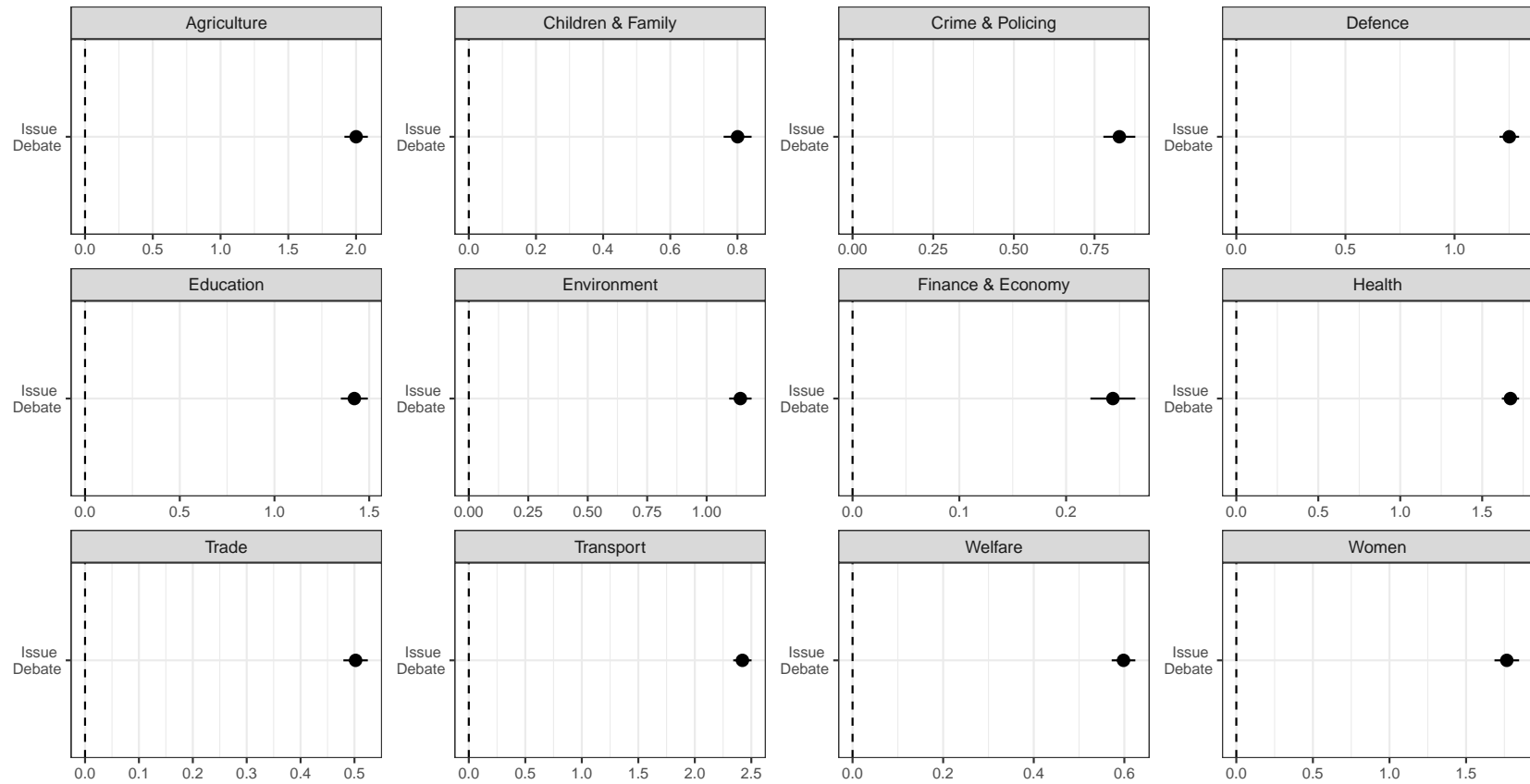


Figure B.2: Average topic use in debates on the same topics

### B.3 Measure correlation

In figure B.3, I show how each of the measures are correlated with one another at the speech-level. Notably, in all cases, the topics are either positively correlated or uncorrelated with one another. This is likely the result of the fact that the word-embedding based measure is picking up on substance overall.

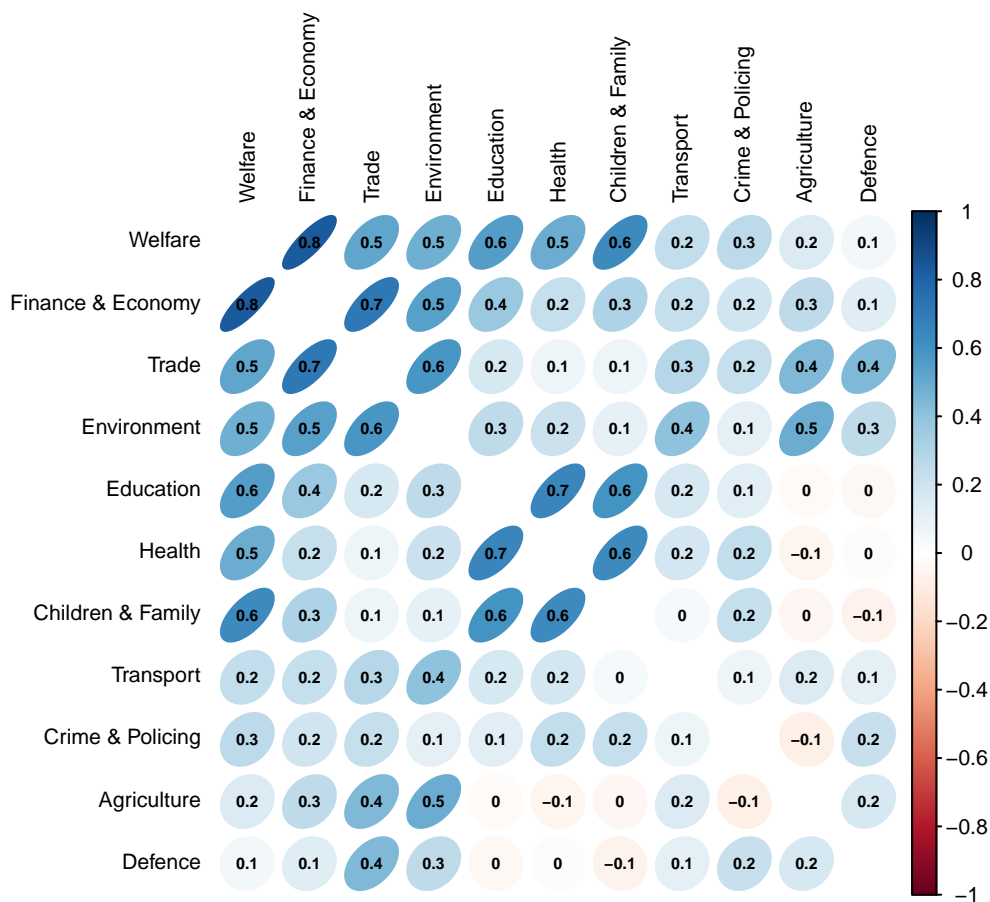


Figure B.3: Correlation between issue areas

## B.4 Structural topic model-based seniority analysis

On which topics do politicians dedicate increased attention the longer they serve in office? As noted in chapter 4, a major puzzle that emerges from the main analysis conducted is that, for almost all issues, I report a downward trajectory for men and women politicians alike. The word-embedding based approach that I use to measure the sets of topics that politicians raise captures only the set of topics that have *a priori* been specified. While these topics capture many of the key issues that dominate British political debate, they do not constitute a complete decomposition of the topical agenda in the UK parliament. As such, in this section, I use an alternative quantitative text technique – structural topic models – that, by design, capture the entire population of topics that MPs raise in the speeches they deliver. In doing so, I can identify the topical areas that senior MPs dedicate their political speeches to raising.

To estimate the full set of topics that MPs raise, I first estimate a correlated topic model (Blei and Lafferty, 2006) (CTM) for all speeches in the data. The CTM is an unsupervised learning approach which assumes that the frequency with which words co-occur within different speeches provides information about the topics that feature in those speeches. The CTM requires the analyst to choose the number of topics,  $K$ , and here I present results from a  $K = 20$  topic model. I implement the CTM as the null form of the Structural Topic Model, which I implement in R (Roberts, Stewart and Tingley, 2014).

The key output of the topic model is  $\theta$ , a  $N * D$  matrix of topic proportions that measures the degree to which each speech ( $i$ ) in the data features each of the estimated topics ( $d$ ).  $\theta_{i,k}$  therefore gives the proportion of speech  $i$  devoted to topic  $d$ . Therefore, unlike the word-embedding based approach, I capture the full set of topics that politicians raise. With these topical proportions in hand, I estimate a series of models for the topic-proportions  $\theta_{i,k}$  of a speech of the following form:

$$\theta_{i,k} = \alpha + \beta_1 \text{Years in Parliament} + \epsilon_{id} \quad (\text{B.3})$$

where  $\alpha$  (i.e., the baseline) represents the average proportion of each topic for an MP with no years of parliamentary experience, and  $\beta_1$  describes the average effect of each additional year in parliament on the prevalence of each topic proportion. My key interest here is identifying the models for which  $\beta_1$  is positive and significant, as this will enable me to identify the sets of topics for which politicians *increase* their attention to the longer they are in parliament. Of course, the time someone is in parliament, and the sets of topics they raise, are likely to be confounded by factors such as political party or ministerial

positions. As such, I also estimate the following:

$$\theta_{i,k} = \alpha + \beta_1 \text{Years in Parliament} + \sum_{k=1}^K \gamma_k X_{i,k} + \epsilon_{i,d} \quad (\text{B.4})$$

where  $X_{i,k}$  is a vector of time-varying MP-level covariates (party, cabinet membership, shadow cabinet membership, government or opposition status, minister, shadow minister, committee chair, gender, and marginality). I present the results from both models in figure B.4, where the coefficients relating to the bivariate specifications are represented by the black solid points and the coefficients relating to the multivariate specifications are represented by the grey hollow points.

The figure presents several key insights. First, it shows that the word-embedding based model clearly did not capture the full scope of topics on the parliamentary agenda, as there are several topics for which politicians increase their attention to with increased parliamentary seniority. Second, and intuitively, the topics which increase in prevalence with parliamentary experience are the topics which are not captured by the word-embedding based approach. In particular, the topics with the largest positive effects seem to relate to procedural topics, such as legislative language, parliamentary address terms, and parliamentary processes. There do also seem to be a handful of substantive topics for which politicians increase their attention the longer they are in parliament: such as one relating to local government and authorities, which is not included in the original set of policy areas I identify. Finally, the topic model has also detected a handful of topics which, despite having a negative relationship with parliamentary experience, were not captured by the word-embedding based approach. For instance, a topic relating to devolution and the devolved assemblies and Brexit. Overall, this analysis shows that, as politicians spend more time in parliament, they seem to dedicate increasing attention to matters of procedure and process relative to junior politicians.

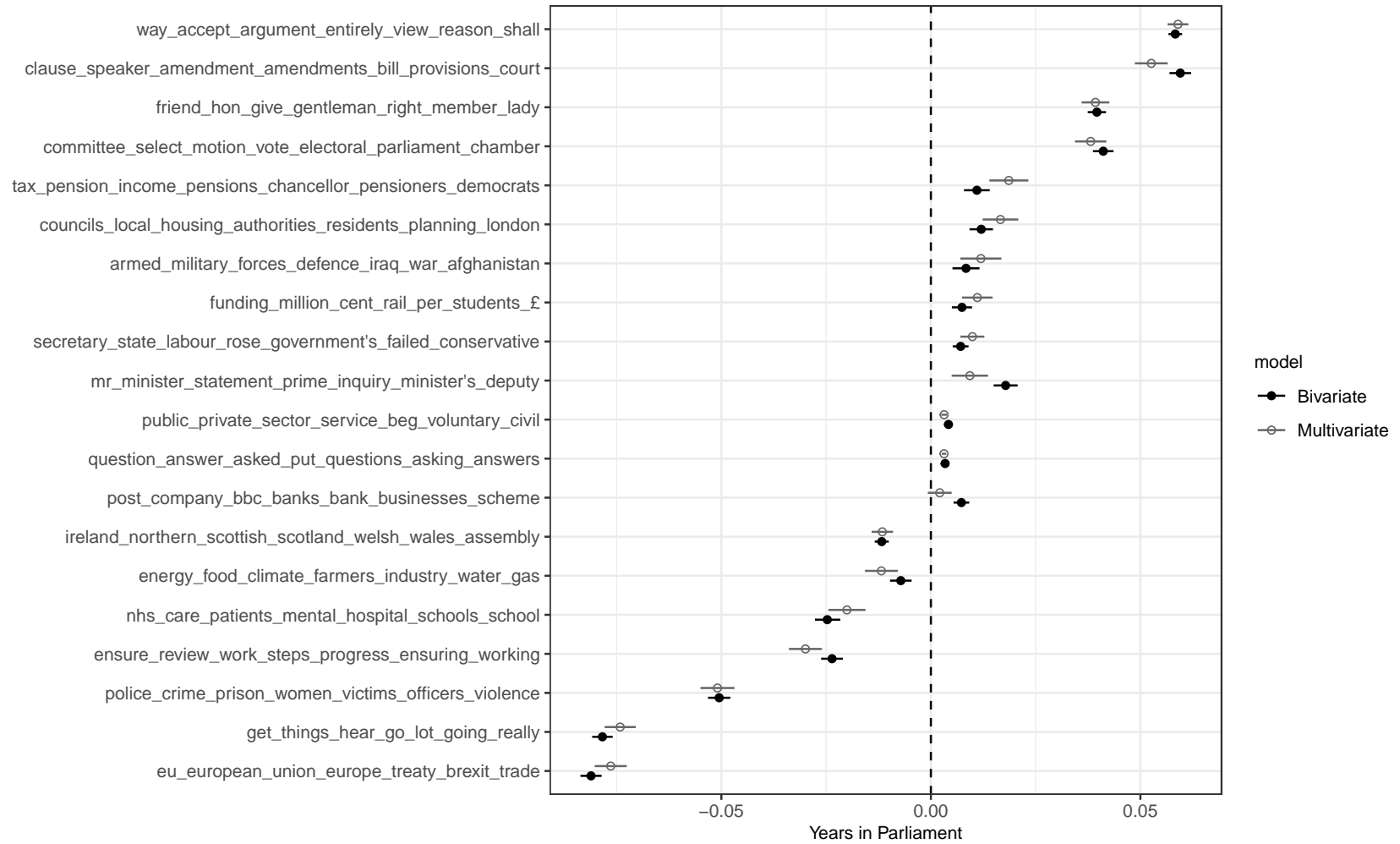


Figure B.4: **Effect of parliamentary experience on issue prevalence:** The figure shows the relationship between years in parliament and the CTM topics,  $K$ .



## B.5 Robustness checks

In tables B.12–B.14, I assess the extent to which the results presented in chapter 4 are robust to alternative specifications. In each of the tables below, I estimate five models for the topic-based outcomes  $Y_{id}^t$  for an individual MP  $i$  in a debate  $d$  for topic  $t$  of the following form:

$$Y_{id}^t = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \epsilon_i \quad (\text{B.5})$$

where  $\beta_1$  describes the difference between men and women with no years of parliamentary experience.  $\beta_2$  describes the effect of increased years in parliament on the extent to which men discuss topics.  $\beta_3$  is the key quantity of interest in this and all specifications, as it describes the difference in the effect of increased time in parliament on the discussion of different topics for men and women. Next, to assess whether the results are robust to MP-specific controls, I estimate the following:

$$Y_{id}^t = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \sum_{k=1}^K \gamma_k X_{ik} + \epsilon_i \quad (\text{B.6})$$

where  $X_{ik}$  is a vector of time-varying MP-level covariates (party, cabinet membership, shadow cabinet measurement, government or opposition, minister, shadow minister, committee chairs, and marginality). Next, to assess whether the results are robust to strict topic-specific controls, I estimate the following:

$$Y_{id}^t = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \sum_{k=1}^K \gamma_k X_{ik} + \epsilon_i \quad (\text{B.7})$$

where  $X_{ik}$  is a vector of generic time-varying MP-level covariates (party, government or opposition, marginality) and a set of dummy variables for government or opposition ministerial positions that relate to each specific topic. These are the same sets of variables described above, which match each

of the issue areas for whether any given MP holds a relevant ministerial role for either the government or opposition in the period of 1997–2019. Fourth, to assess whether the results are robust when controlling for *within*-MP change over the course of their careers, I estimate the following:

$$Y_{id}^t = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \sum_{k=1}^K \gamma_k X_{ik} + \lambda_i + \epsilon_i \quad (\text{B.8})$$

where  $X_{ik}$  represents a set of time-varying MP-level covariates (cabinet membership, shadow cabinet measurement, government or opposition, minister, shadow minister, committee chairs, and marginality) and  $\lambda_i$  represents MP fixed effects. Finally, it seems likely that the prevalence of topics may change over the time period under study. To assess whether the results are robust to factors which change over time for all politicians, I estimate the following:

$$Y_{idt}^t = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \sum_{k=1}^K \gamma_k X_{ik} + \lambda_i + \delta_t + \epsilon_i \quad (\text{B.9})$$

where  $X_{ik}$  represents a set of time-varying MP-level covariates (cabinet membership, shadow cabinet measurement, government or opposition, minister, shadow minister, committee chairs, and marginality),  $\lambda_i$  represents MP fixed effects, and  $\delta_t$  represents fixed effects for parliamentary terms of which there are 6 between 1997 and 2019. Throughout all five model specifications, my quantity of interest is  $\beta_3$ . In all five model specifications, I cluster standard errors at the MP level  $i$ . By increasing how conservative the specifications are in each of the five sets of models, I can observe how robust my central argument – that while junior women parliamentarians who lack experience in the eyes of the public and their colleagues may feel pressured to talk only about a narrow set of issues stereotypical associated with women, this pressure diminishes as seniority increases – is to alternative explanations. I find that, for all policy areas, the results are robust to these alternative model specifications.

Table B.11: Relationship between MP gender and years in parliament on agriculture topic

	<i>Dependent variable:</i>				
	Agriculture				
	(1)	(2)	(3)	(4)	(5)
Woman MP	-0.022 (0.023)	-0.030 (0.023)	-0.028 (0.023)		
Years in Parliament	-0.009*** (0.001)	-0.011*** (0.002)	-0.011*** (0.002)	-0.006*** (0.001)	0.003 (0.002)
Woman MP*Years in Parliament	-0.002 (0.003)	-0.001 (0.003)	-0.002 (0.003)	0.001 (0.003)	0.0003 (0.003)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.006	0.010	0.012	0.001	0.001
Adjusted R <sup>2</sup>	0.006	0.010	0.012	-0.003	-0.002

Note: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

Table B.12: Relationship between MP gender and years in parliament on children and family topic

	<i>Dependent variable:</i>				
	Children and Family				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.454*** (0.033)	0.432*** (0.030)	0.456*** (0.033)		
Years in Parliament	-0.011*** (0.002)	-0.009*** (0.002)	-0.011*** (0.002)	-0.003** (0.001)	0.009*** (0.002)
Woman MP*Years in Parliament	-0.015*** (0.005)	-0.015*** (0.004)	-0.015*** (0.005)	-0.010*** (0.004)	-0.010*** (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,503	437,503	437,503
R <sup>2</sup>	0.040	0.052	0.040	0.001	0.001
Adjusted R <sup>2</sup>	0.039	0.052	0.040	-0.003	-0.002

Note: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

Table B.13: Relationship between MP gender and years in parliament on crime and policing topic

	<i>Dependent variable:</i>				
	Crime and Policing				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.143*** (0.024)	0.138*** (0.023)	0.144*** (0.023)		
Years in Parliament	-0.003** (0.001)	-0.002 (0.002)	-0.002 (0.002)	-0.003*** (0.001)	-0.0004 (0.002)
Woman MP*Years in Parliament	-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.002)	-0.005* (0.003)	-0.005* (0.003)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.004	0.009	0.009	0.0004	0.001
Adjusted R <sup>2</sup>	0.004	0.009	0.009	-0.003	-0.002

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.14: Relationship between MP gender and years in parliament on defence topic

	<i>Dependent variable:</i>				
	Defence				
	(1)	(2)	(3)	(4)	(5)
Woman MP	-0.051* (0.028)	-0.054** (0.027)	-0.048* (0.028)		
Years in Parliament	-0.001 (0.002)	-0.001 (0.003)	-0.003 (0.003)	0.007*** (0.002)	0.002 (0.002)
Woman MP*Years in Parliament	-0.004 (0.005)	-0.004 (0.004)	-0.005 (0.004)	-0.007** (0.003)	-0.007** (0.003)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.002	0.007	0.012	0.001	0.002
Adjusted R <sup>2</sup>	0.002	0.007	0.012	-0.002	-0.001

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.15: Relationship between MP gender and years in parliament on education topic

	<i>Dependent variable:</i>				
	Education				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.294*** (0.031)	0.284*** (0.027)	0.286*** (0.027)		
Years in Parliament	-0.013*** (0.002)	-0.012*** (0.002)	-0.012*** (0.002)	-0.006*** (0.002)	-0.0005 (0.002)
Woman MP*Years in Parliament	-0.012*** (0.004)	-0.013*** (0.003)	-0.013*** (0.003)	-0.016*** (0.004)	-0.015*** (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.028	0.049	0.050	0.002	0.003
Adjusted R <sup>2</sup>	0.028	0.049	0.050	-0.001	-0.0003

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.16: Relationship between MP gender and years in parliament on environment topic

	<i>Dependent variable:</i>				
	Environment				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.118*** (0.031)	0.108*** (0.028)	0.114*** (0.029)		
Years in Parliament	-0.016*** (0.002)	-0.016*** (0.002)	-0.016*** (0.002)	-0.005*** (0.002)	-0.006*** (0.002)
Woman MP*Years in Parliament	-0.006 (0.005)	-0.007 (0.004)	-0.007* (0.004)	-0.009** (0.004)	-0.008* (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.024	0.043	0.041	0.001	0.002
Adjusted R <sup>2</sup>	0.024	0.043	0.041	-0.002	-0.002

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.17: Relationship between MP gender and years in parliament on economy topic

	<i>Dependent variable:</i>				
	Finance and Economy				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.074*** (0.028)	0.056** (0.026)	0.065** (0.027)		
Years in Parliament	-0.014*** (0.002)	-0.013*** (0.002)	-0.013*** (0.003)	-0.009*** (0.002)	-0.011*** (0.002)
Woman MP*Years in Parliament	-0.006 (0.004)	-0.006** (0.003)	-0.006* (0.003)	-0.009** (0.004)	-0.007** (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.017	0.032	0.029	0.002	0.003
Adjusted R <sup>2</sup>	0.017	0.032	0.029	-0.001	-0.0001

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.18: Relationship between MP gender and years in parliament on health topic

	Dependent variable:				
	Health				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.381*** (0.033)	0.364*** (0.031)	0.373*** (0.031)		
Years in Parliament	-0.012*** (0.002)	-0.012*** (0.002)	-0.013*** (0.002)	-0.006*** (0.001)	0.003 (0.002)
Woman MP*Years in Parliament	-0.013*** (0.005)	-0.013*** (0.004)	-0.014*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.034	0.050	0.054	0.002	0.002
Adjusted R <sup>2</sup>	0.034	0.050	0.054	-0.002	-0.001

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.19: Relationship between MP gender and years in parliament on trade topic

	Dependent variable:				
	Trade				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.010 (0.029)	0.005 (0.028)	0.012 (0.029)		
Years in Parliament	-0.011*** (0.002)	-0.011*** (0.003)	-0.010*** (0.003)	0.002 (0.002)	-0.006*** (0.002)
Woman MP*Years in Parliament	-0.004 (0.005)	-0.004 (0.004)	-0.005 (0.004)	-0.009** (0.004)	-0.010*** (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.010	0.020	0.017	0.0002	0.001
Adjusted R <sup>2</sup>	0.010	0.020	0.017	-0.003	-0.002

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.20: Relationship between MP gender and years in parliament on transport topic

	Dependent variable:				
	Transport				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.029 (0.025)	0.020 (0.023)	0.024 (0.023)		
Years in Parliament	-0.011*** (0.001)	-0.011*** (0.002)	-0.012*** (0.002)	-0.004*** (0.001)	-0.004** (0.002)
Woman MP*Years in Parliament	-0.002 (0.004)	-0.002 (0.003)	-0.002 (0.003)	-0.005 (0.003)	-0.005 (0.003)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.010	0.022	0.026	0.001	0.001
Adjusted R <sup>2</sup>	0.010	0.021	0.026	-0.003	-0.002

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.21: Relationship between MP gender and years in parliament on welfare topic

	Dependent variable:				
	Welfare				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.292*** (0.034)	0.269*** (0.030)	0.277*** (0.031)		
Years in Parliament	-0.019*** (0.003)	-0.017*** (0.003)	-0.017*** (0.003)	-0.010*** (0.002)	-0.004** (0.002)
Woman MP*Years in Parliament	-0.012** (0.006)	-0.013*** (0.004)	-0.013*** (0.005)	-0.012*** (0.004)	-0.011*** (0.004)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.045	0.067	0.065	0.004	0.005
Adjusted R <sup>2</sup>	0.045	0.067	0.065	0.0005	0.002

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table B.22: Relationship between MP gender and years in parliament on references to women

	Dependent variable:				
	Women				
	(1)	(2)	(3)	(4)	(5)
Woman MP	0.184*** (0.016)	0.193*** (0.015)	0.190*** (0.015)		
Years in Parliament	0.0004 (0.001)	0.002* (0.001)	0.001 (0.001)	0.004*** (0.001)	-0.007*** (0.002)
Woman MP*Years in Parliament	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	0.005* (0.003)	0.003 (0.003)
Controls?	No	Yes	Yes	Yes	Yes
MP FEs?	No	No	No	Yes	Yes
Parliamentary term FEs?	No	No	No	No	Yes
Observations	437,503	437,500	437,500	437,503	437,503
R <sup>2</sup>	0.005	0.008	0.007	0.001	0.003
Adjusted R <sup>2</sup>	0.005	0.008	0.007	-0.003	-0.001

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

## B.6 Controlling for individual-level covariates

In this section, I present the estimates for each of the individual-level covariates  $\gamma X_{ik}$  for each policy area described in chapter 4, where  $\gamma X_{ik}$  includes:

- Party (categorical: Conservative; Labour; Liberal Democrat; Green; Other)
- Years in parliament (in years, continuous)
- Attends cabinet (binary)
- Attends shadow cabinet (binary)
- Government minister (binary)
- Opposition minister (binary)
- Committee chair (binary)
- Margin of victory in prior election (percentage points, continuous)
- Government MP (binary)

I present the estimates for each of the individual-level covariates in figure B.5, where the point estimates and confidence are coloured in black if the difference is associated with a  $p$ -value of less than 0.05, and otherwise in grey. Although not my primary quantity of interest, there are several interesting relationships to note. First, I see intuitive partisan differences. Labour MPs are significantly more likely to talk about social policy issues than Conservative MPs. Reassuringly, I also see a large positive effect for the Greens for the environment. This provides further face validity that my word-embedding based measures are detecting the appropriate topics. Second, in many cases the size of the coefficients for the difference between men and women MPs are larger than many of the other individual-level covariates. At the aggregate level, therefore, MP gender seems to be a strong predictor in determining the sets of issues politicians raise. Overall, the patterns here show interesting differences in the characteristics that determine the policy areas that MPs raise, and further reassurance that the measures are working as intended.



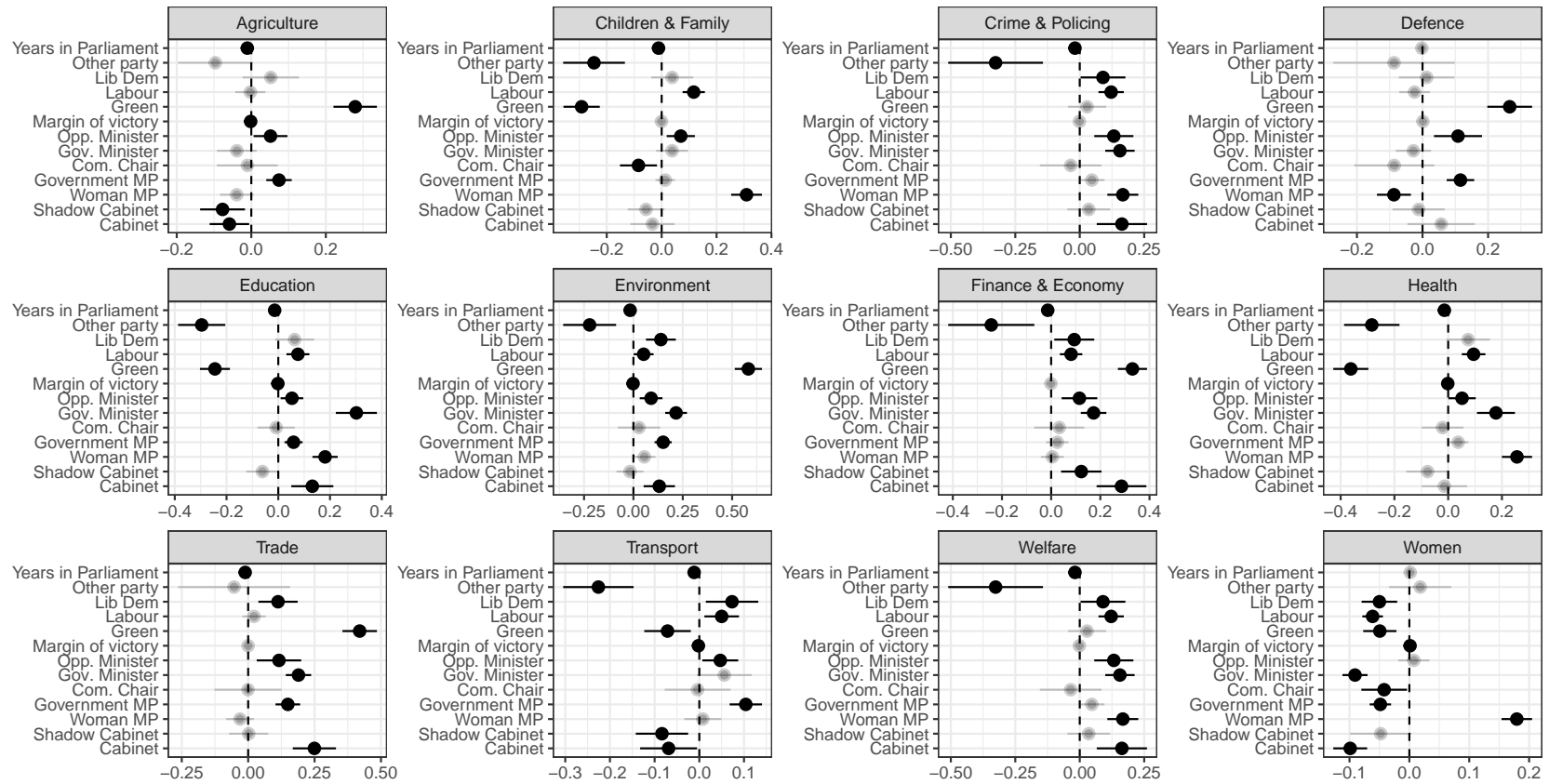


Figure B.5: Individual-level covariate effects

## B.7 Raising women’s perspectives across issue areas

In chapter 4, I present the results from an analysis where I identify whether women politicians raise women’s perspectives more than men in debates on all policy areas. I show that this is the case for all policy areas. In this section, I also assess whether these differences are consistent with increased seniority. To do so, I subset to debates on each of the topics and estimate the following:

$$Y_{id} = \alpha + \beta_1 \text{WomanMP}_i + \beta_2 \text{YearsinParliament}_i + \beta_3 (\text{WomanMP}_i \cdot \text{YearsinParliament}_i) + \sum_{k=1}^K \gamma_k X_{ik} + \epsilon_i \quad (\text{B.10})$$

where  $\beta_1$  describes the difference in the extent to which men and women raise women’s perspectives with no years of parliamentary experience.  $\beta_2$  describes the effect of increased years in parliament on the extent to which men talk about women.  $\beta_3$  describes the difference in the effect of increased time in parliament on raising women’s perspectives between men and women.  $X_{ik}$  is a vector of time-varying MP-level covariates (party, cabinet membership, shadow cabinet measurement, government or opposition, minister, shadow minister, committee chairs, and marginality). I cluster standard errors at the individual MP level  $i$ .

The results for each topic area are presented in tables B.23–B.24. As is clear from all models, the coefficients for  $\beta_1$  are all positive and significant, suggesting that, among junior politicians, women are significantly more likely to raise women’s perspectives than men. However, also clear from the tables is that this difference does not seem to change because of increased parliamentary experience. The coefficients for  $\beta_3$  are extremely small and insignificant in all models except for transport debates. The implication of these findings is positive, as they suggest that across debates on a wide variety of topics, women politicians raise women’s perspectives, and they continue to do this even with increased parliamentary experience.

Table B.23: Relationship between MP gender and years in parliament on amplifying women's perspectives across debate types

	<i>Dependent variable:</i>					
	Family (1)	Health (2)	Education (3)	Welfare (4)	Crime (5)	Transport (6)
Woman MP	0.159*** (0.042)	0.080*** (0.023)	0.152*** (0.025)	0.166*** (0.020)	0.213*** (0.023)	0.121*** (0.024)
Years in Parliament	-0.001 (0.002)	0.00004 (0.001)	-0.002* (0.001)	-0.003*** (0.001)	0.001 (0.001)	-0.0005 (0.001)
Woman MP*Years in Parliament	-0.001 (0.004)	0.002 (0.002)	-0.004 (0.003)	-0.0004 (0.002)	0.002 (0.002)	-0.005** (0.002)
Observations	8,570	19,332	22,473	32,127	19,614	12,469
R <sup>2</sup>	0.009	0.004	0.008	0.009	0.014	0.004
Adjusted R <sup>2</sup>	0.007	0.003	0.007	0.008	0.014	0.003

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table B.24: Relationship between MP gender and years in parliament on amplifying women's perspectives across debate types

	<i>Dependent variable:</i>				
	Environment (1)	Economy (2)	Trade (3)	Agriculture (4)	Defence (5)
Woman MP	0.161*** (0.029)	0.131*** (0.013)	0.204*** (0.024)	0.083** (0.037)	0.170*** (0.018)
Years in Parliament	-0.004*** (0.001)	0.002*** (0.0005)	-0.0004 (0.001)	-0.002 (0.001)	-0.001** (0.001)
Woman MP*Years in Parliament	-0.005* (0.003)	0.001 (0.001)	0.0001 (0.002)	-0.002 (0.004)	-0.002 (0.002)
Observations	19,252	69,621	32,286	10,271	30,564
R <sup>2</sup>	0.009	0.008	0.011	0.007	0.008
Adjusted R <sup>2</sup>	0.008	0.008	0.010	0.006	0.007

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## B.8 Gender differences in policy areas over time

While not the primary substantive quantity of interest in chapter 4, gender differences in the traditional divisions in the sets of issues politicians may raise may also have changed over time. Several factors in the UK over the past 25 years are likely to have decreased the degree to which UK MPs, and especially women, need to conform to stereotype-expected behaviours. Changes to the social roles played by women in public life, and in politics, may have reduced the validity and content of stereotypes in the eyes of the public. As such, voters may now have less clear preferences in the sets of issues they expect men and women to have expertise in, and, may now be less likely to sanction women politicians for their behaviour. I here assess whether gender differences in the issues men and women raise may have changed between 1997 and 2019 by estimating models for the topic-based outcomes  $Y_{id}^t$  for an individual MP  $i$  in a debate  $d$  for topic  $t$  of the following form:

$$\begin{aligned}
 Y_{id}^t = & \alpha + \beta_1 \text{WomanMP}_i + \sum_{l=2}^L \varphi_l^1 \text{Year}_{i,l} + \\
 & \sum_{l=2}^L \varphi_l^2 (\text{WomanMP}_i \cdot \text{Year}_{i,l}) + \\
 & \sum_{k=1}^K \gamma_k X_{i,k} + \epsilon_{idt}
 \end{aligned} \tag{B.11}$$

where  $\text{Year}$  is a matrix of dummy variables with cells which take the value of 1 when MP  $i$  is delivering a speech in  $l$  year in the data.  $\varphi_l^1$  therefore captures the effect of spending  $l$  additional time in office (relative to an MP being in parliament in 1997, the first year in the data) on the attention given to topic  $t$  for men MPs, and  $\varphi_l^2$  describes the difference in this effect for women MPs. Accordingly,  $\varphi_l^1 + \varphi_l^2$  describes the effect of each additional year  $l$  (compared to 1997) on the outcome for women MPs.  $X_{ik}$  is a vector of time-varying MP-level covariates (party, age, cabinet membership, shadow cabinet membership, minister, shadow minister, years in parliament, marginality, and government or opposition status). As there are multiple observations per MP, I cluster standard errors at the MP level  $i$ . I present the results of this analysis in figure B.6.

The figure shows that while there are some year-by-year changes in the gender gaps for some of the topic areas, the differences tend to follow no particularly clear patterns. For the traditionally “feminine” policy areas, I see that women MPs talk consistently more about these issues than men, and that the gender gap is statistically significant in all years in the data. The figure therefore shows while the gender differences shift slightly over time, there is no clear trend that women talk less about

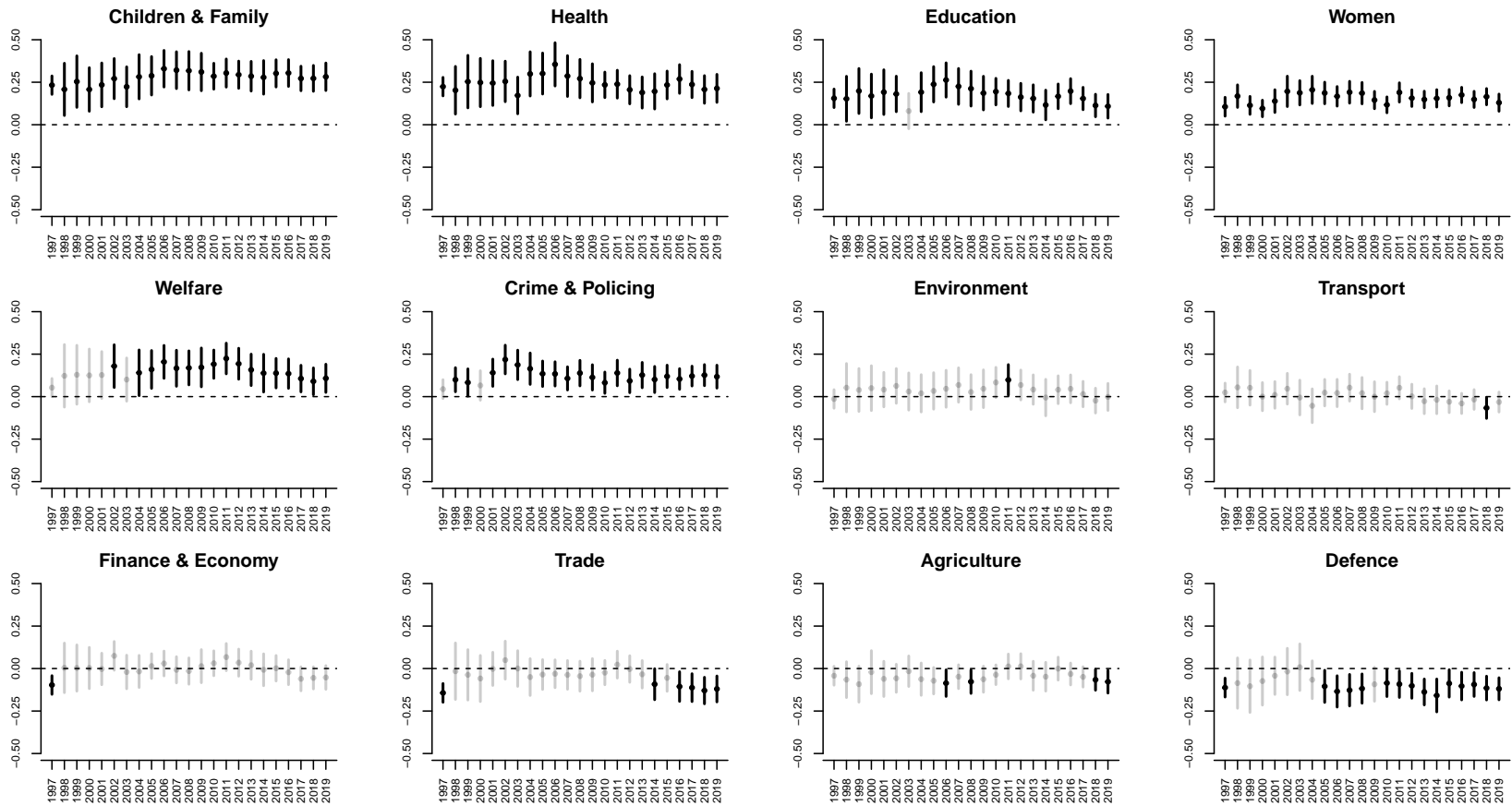


Figure B.6: Gender differences in issues over time

traditional “feminine”, nor more about traditional “masculine”, issue areas over time. In figure B.7, I depict the change in topic prevalence over time separately for men and women. The figure shows that, broadly speaking, men and women follow very similar trajectories in the topics they focus on over time.

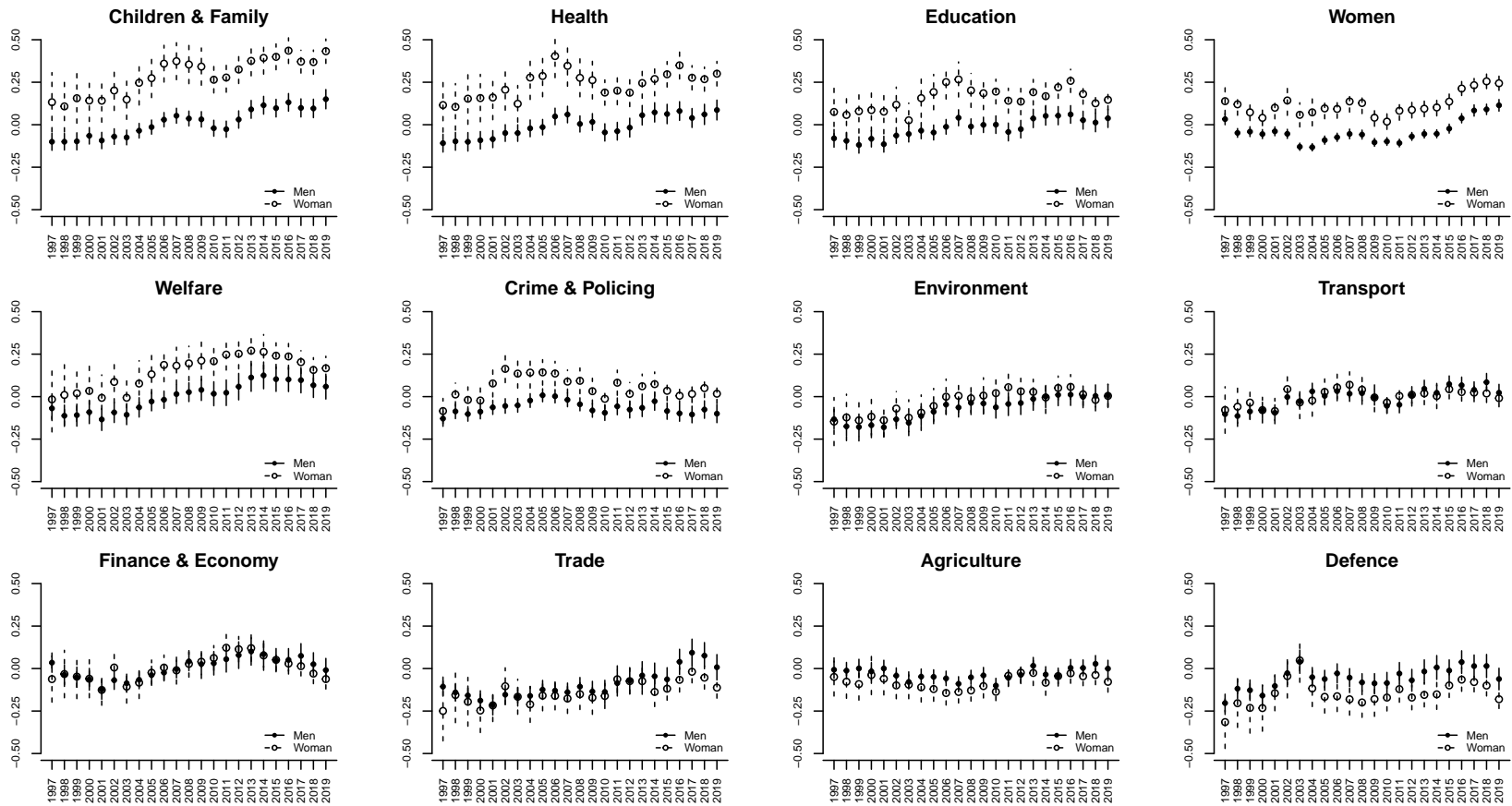


Figure B.7: Average topic use for men and women over time

# Appendix C

## Appendix C – A Double Standard? Gender Bias in Voters’ Perceptions of Political Arguments

### C.1 Experiment pre-test survey

Prior to fielding the survey experiment, a pre-test survey was fielded by Prolific to 1,499 members of their UK online panel in August 2021. Speeches were pre-tested without the MP’s gender stated, and there were 3 x 2 x 2 (style x treatment status x policy area) = 18 speeches tested overall. Following an introduction screen describing the task, respondents were randomly assigned to read one speech and asked to answer several questions about the speech and the MP. First, depending on the style assigned, respondents were asked whether they agreed or disagreed that the speech was “emotional”, “aggressive” or “evidence-based”. The purpose of this was to ensure that the treatments were representative of the styles, and that the controls were not representative of the styles. An example of the task and question can be seen in figure C.1. Results from the style pre-testing exercise are presented in table C.1. Responses are presented as the extent to which respondents (strongly) disagree, neither agree nor disagree, or (strongly) agree that the speech was “emotional”, “aggressive” or “evidence-based”. Figures C.2–C.4 show the responses collapsed by style, style treatment status, and policy area.

Table C.1: Pre-testing results

Style	Treatment status	(Strongly) disagree	Neither	(Strongly) agree
Emotion	Control	60%	21.7%	18.3%
Emotion	Treatment	10.1%	17.4%	72.5%
Aggression	Control	65.2%	10.4%	24.4%
Aggression	Treatment	19.9%	12.4%	67.7%
Evidence	Statistics	6.4%	12.8%	80.8%
Evidence	Anecdote	28.8%	14.4%	56.8%



"I have to start by saying I utterly disagree with what you've just said. You clearly lack any understanding of how serious this is, and, frankly, you show absolutely no care for the people we represent. The way we understand diseases and how to treat them has grown over the years. There has been significant innovation in treatments that might help save the lives of people in this country and we should be committed to using these cutting-edge treatments and therapies to help save and improve patients' lives. But often nothing is done because it is simply deemed not worth the money. I am utterly revolted by the idea that people are not getting treated merely because of some appalling cost-benefit calculation. This is disgusting, deplorable, and inhuman. It is so important that how we fund and pay for drugs works for everybody. Things must change, and anyone who opposes this cannot claim they care about the well-being of our people."

Do you agree or disagree that the argument made by the MP is aggressive?

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly agree

Figure C.1: Example of pre-testing task (Aggression, treatment, health)

Overall, the results from the style pre-testing exercise confirm that the treatments are perceived as representative of the styles. For instance, 72.5% of respondents assigned to the emotion treatment conditions either agreed or strongly agreed that they were emotional. Similarly, 67.7% of respondents assigned the aggression treatment conditions agree or strongly agreed that they were aggressive. Further, the pre-testing results provide good evidence that the control texts were not perceived to be particularly representative of the styles. For instance, 60% of respondents assigned to the emotion control conditions either disagreed or strongly disagreed that they were emotional. Additionally, 65.2% of the respondents assigned to the aggression control conditions either disagreed or strongly disagreed that they were aggressive. While most respondents perceived that both the statistics and anecdote speeches were evidence-based, a larger percentage of respondents perceived the statistical speeches as evidence based than the anecdotal speeches (80.8% and 56.8% respectively). Further, the results are consistent across the policy areas for each style. For example, for the emotion controls, between

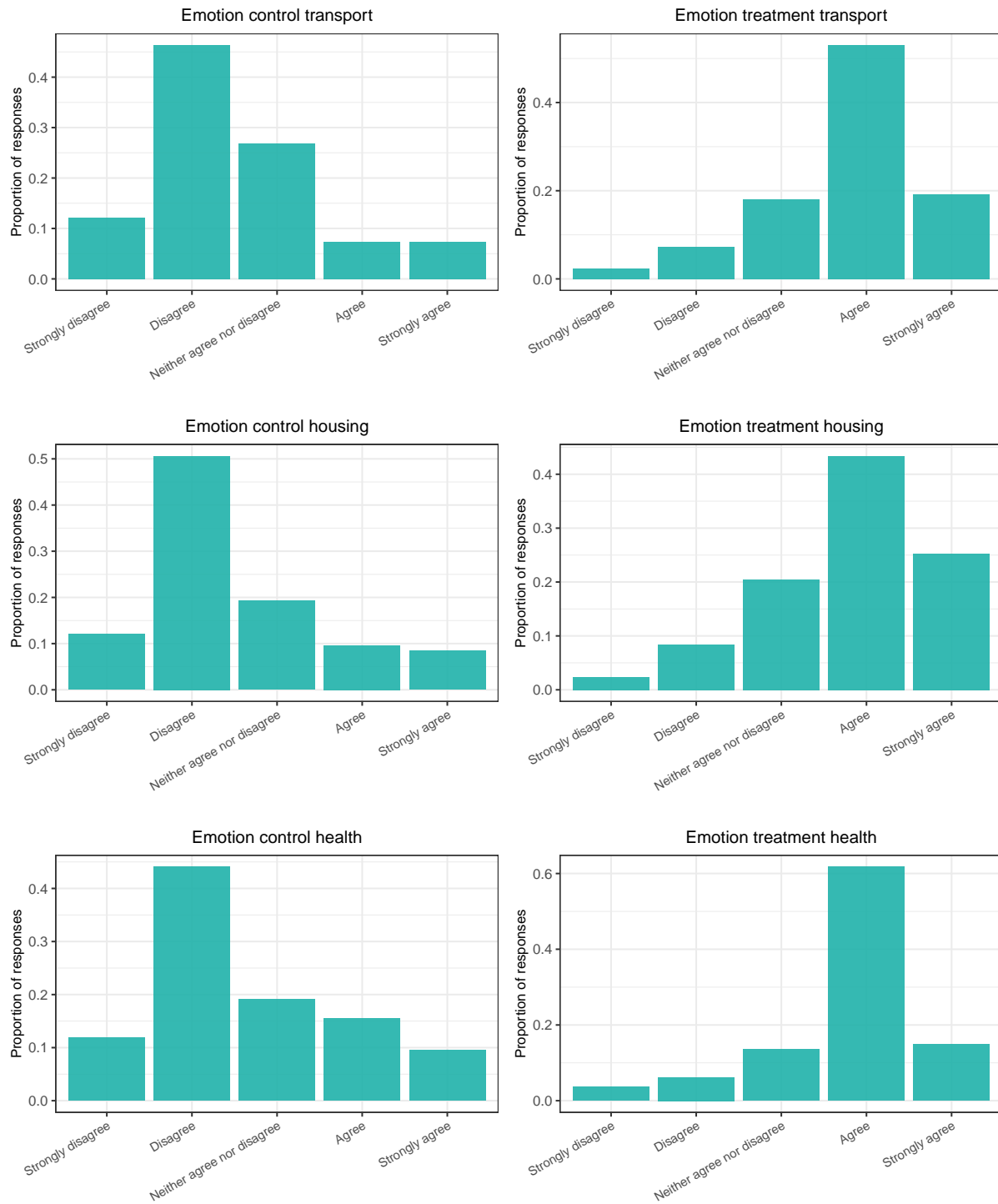


Figure C.2: Distribution of responses across treatment group status for emotion

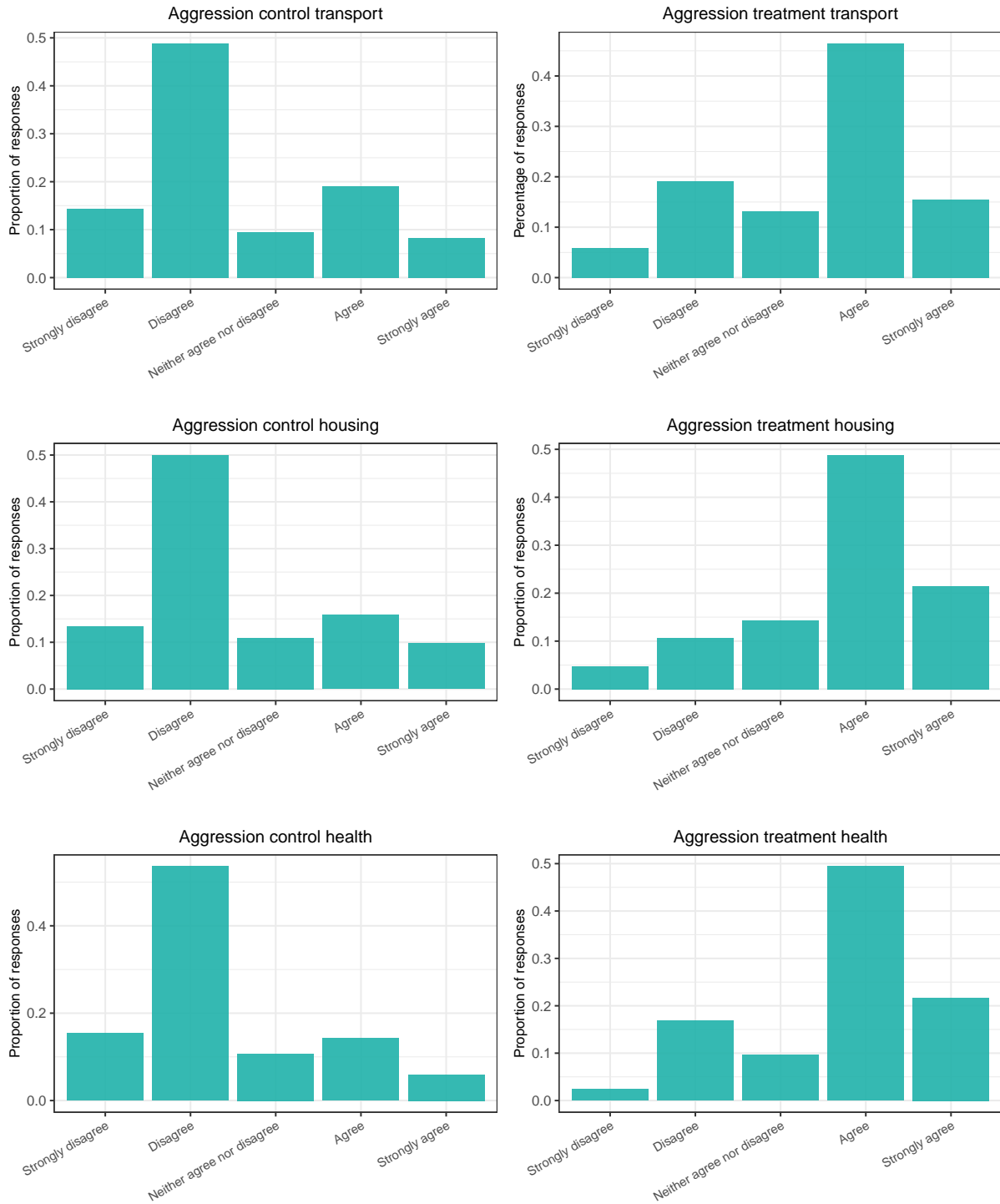


Figure C.3: Distribution of responses across treatment group status for aggression

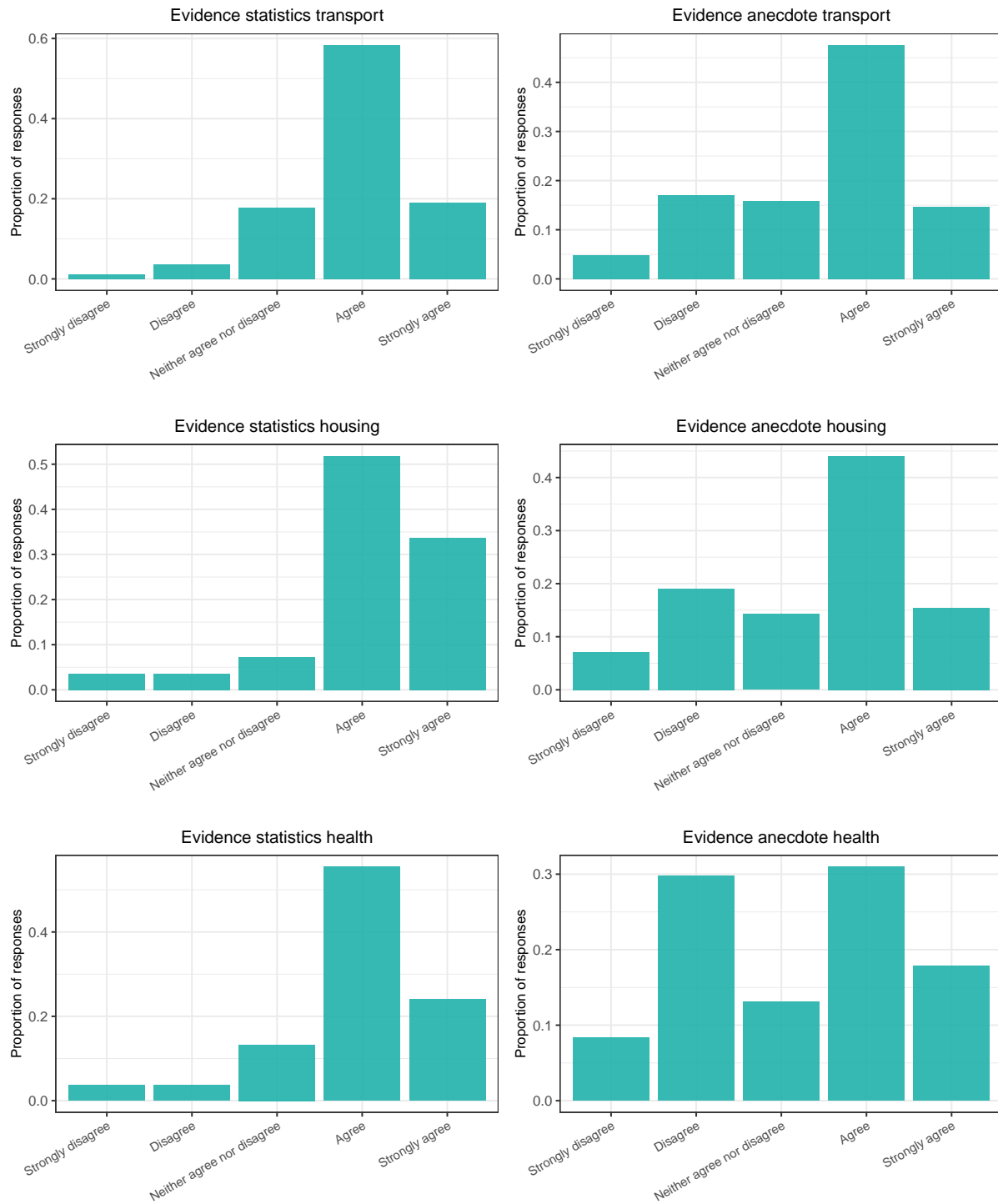


Figure C.4: Distribution of responses across treatment group status for evidence

45%–50% of respondents “disagreed” that the arguments were emotional across the three policy areas, and roughly 10% of respondents “strongly disagreed” that the arguments were emotional across the three policy areas. Similarly, for the aggression treatments between roughly 45%–50% of respondents “agreed” that the arguments were aggressive.

Second, to ensure that treatment texts were not too long for respondents to read and understand, all respondents were asked the follow-up question “Did you find this text too long and/or too complicated to understand?”. Overall, across all 18 treatments, 95.7% said that the speeches were not too long and/or complicated to understand, and only 4.3% said they were. Overall, there is good evidence that the style treatments work as planned, and that the treatment texts are not too long for respondents to read and understand.

## C.2 Sample recruitment

The participants for the survey experiment were voluntarily recruited from YouGov's UK online panel of respondents. YouGov recruits panel members "from a host of different sources, including via standard advertising, and strategic partnerships with a broad range of websites".<sup>1</sup> Panel members are awarded points for completing YouGov surveys which they can later redeem for financial reward. As such, they are compensated financially for their participation and this is handled by YouGov directly. For the survey YouGov draws a sub-sample of the panel that is representative of British adults in terms of age, gender, social class and education, and invites this sub-sample to complete the survey. Demographic information of panel members is collected by YouGov when members of the public sign up to be participants for the panel. Therefore, this data is already held by YouGov and was not actively collected for this project.

The survey experiment was fielded as a wider part of the YouGov UK Political Team's Omnibus, of which all respondents consent to participate in. Participants were made aware that their answers remained anonymous and would be used for research purposes only. Further, as part of consenting to participate, they are made aware that they can leave the survey at any time, however once they complete the survey their answers could not be retracted. Participants were also aware that I would not receive any information from YouGov that would enable me to identify them. In the design I describe in chapter 5, respondents are presented with hypothetical politicians and arguments that were written by the researcher. Therefore, so as not to deceive participants, this was made clear through the following statements prior to and after completing the survey experiment:

**Opening statement:** *A central part of a Member of Parliament's job is to debate political issues. MPs take many different approaches in how they argue in debate. We would like to present you with an argument that mimics the typical kind of speech that MPs deliver. Please read the argument carefully. Afterwards, we will ask you some questions about the argument and the politician. We will ask you to complete this task three times in total.*

**Closing statement:** *The arguments you were presented were not representative of actual arguments delivered by MPs, and to the best of our knowledge do not reflect any argument delivered by an existing MP.*

---

<sup>1</sup>See "Panel Methodology", YouGov.

### C.3 Treatment texts

The full set of gendered names used in the experiment are displayed in table C.2.

Table C.2: Gendered names for experimental design

Man forename	Woman forename	Surname
Adam	Beth	Craddock
Jack	Charlotte	Jones
Peter	Lucy	Richards

In table C.3, I present the full text of each of the speeches. The “treatment” for each of the styles are indicated in bold font and square brackets.

Table C.3: Treatment texts

Treatment number	Treatment text
Treatment 1: Emotion, Control, Housing	Our housing market has not been as it should for years. It is the job of all of us to increase the supply of housing available. But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. Many people live in unstable rented housing who may be driven out by increasing rent costs. Work should be done to help people move out of rented accommodation and become homeowners. We need different strategies to help to increase supply and make homes more affordable. We should build new homes, and we should repurpose empty homes. Builders, investors, and local councils will need to work together for change to occur. Our housing market does not work for many people. There is a need for new policy to help home ownership become realistic for young people all around the country.
Treatment 2: Emotion, Treatment, Housing	<b>[The idea of owning your own home is one filled with such excitement for people.]</b> It is the job of all of us to increase the supply of housing. <b>[I recall with such warmth that sense of euphoria, exhilaration, and real achievement when I held the keys to my first home. I can, hand on my heart, say that it was one of the proudest days of my life.]</b> But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. We need different strategies to help to increase supply and make homes more affordable. <b>[I am hopeful that, if we work together, we can help people get closer to experiencing the joy of homeownership themselves. Changing the lives of the people I represent to help reach brighter futures is, after all, really what makes my job such a pleasure to do.]</b>
Treatment 3: Aggression, Control, Housing	Our housing market has not been as it should for years. It is the job of all of us to increase the supply of housing available. But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. Many people live in unstable rented housing who may be driven out by increasing rent costs. Work should be done to help people move out of rented accommodation and become homeowners. We need different strategies to help to increase supply and make homes more affordable. We should build new homes, and we should repurpose empty homes. Builders, investors, and local councils will need to work together for change to occur. Our housing market does not work for many people. There is a need for new policy to help home ownership become realistic for young people all around the country.
Treatment 4: Aggression, Treatment, Housing	<b>[I have to start by saying I utterly disagree with what you've just said. You clearly lack any understanding of how serious this is, and, frankly, you show absolutely no care for the people we represent.]</b> It is the job of all of us to increase the supply of housing. But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. <b>[Too many people live in insecure rented houses owned by exploitative, greedy, penny-pinching landlords who milk their tenants for all they're worth.]</b> We need different strategies to help to increase supply and make homes more affordable. <b>[Young people have been utterly abandoned and left fearful about their futures because we have catastrophically failed to do enough to help them. Things must change, and anyone who opposes this cannot claim they care about the well-being of our people.]</b>
Treatment 5: Evidence, Statistics, Housing	Our housing market has not been as it should for years. It is the job of all of us to increase the supply of housing. But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. <b>[Those in their mid-30s to mid-40s are three times more likely to be renters than 20 years ago. People in the early 1990s could expect to pay 3.5 times their annual earnings on buying a home, but this has risen to 7.8 today. In 2019, the average property sold for £235,300, meanwhile the average pay came in at £29,000.]</b> Our housing market does not work for many people. We need different strategies to help to increase supply and make homes more affordable. There is a need for new policy to help homeownership become realistic for young people all around the country.



Treatment number	Treatment text
Treatment 6: Evidence, Anecdote, Housing	Our housing market has not been as it should for years. It is the job of all of us to increase the supply of housing. But for many young people the gap between wages and house prices is too wide for homeownership to be viable any time soon. <b>[I spoke recently to Eleanor and Michael, a young couple in my constituency, who have been renting together for a few years and want to buy their first home. They told me they are extremely concerned that their dream of homeownership simply will not be one they can reach unless something changes.]</b> Our housing market does not work for many people. We need different strategies to help to increase supply and make homes more affordable. There is a need for new policy to help homeownership become realistic for young people, <b>[just like Eleanor and Michael]</b> , all around the country.
Treatment 7: Emotion, Control, Health	The way we understand diseases and how to treat them has grown over the years. There have been advances in treatments that might help save the lives of people in this country and we should be committed to using them. But in some cases, patients are left without access to the drugs they need because of the costs of treatment. Patients being able to access the treatments they require is one of the principles of providing universal health care to everyone who needs it. I think some people find it challenging to accept that questions of money should enter decisions about health care, but this is the situation that many hospitals, doctors, and patients find themselves in. It is important that how we fund and pay for drugs works for everybody. Change should happen to make sure that both medical services and patients get the treatments and resources they need.
Treatment 8: Emotion, Treatment, Health	The way we understand diseases and how to treat them has grown over the years. There have been advances in treatments that might help save the lives of people in this country and we should be committed to using them. <b>[These advances are miraculous. I am brought to tears when I hear heartening stories about young children's lives being saved and getting to see the joy and relief of families getting to spend more time with their loved ones. Being with the ones we love is really the only thing that matters.]</b> But in some cases, patients are left without access to the drugs they need because of the costs of treatment. It is important that how we fund and pay for drugs works for everybody. <b>[This isn't only about improving our country for the better, but about ensuring that we can be hopeful about a future of humanity that shines bright.]</b>
Treatment 9: Aggression, Control, Health	The way we understand diseases and how to treat them has grown over the years. There have been advances in treatments that might help save the lives of people in this country and we should be committed to using them. But in some cases, patients are left without access to the drugs they need because of the costs of treatment. Patients being able to access the treatments they require is one of the principles of providing universal health care to everyone who needs it. I think some people find it challenging to accept that questions of money should enter decisions about health care, but this is the situation that many hospitals, doctors, and patients find themselves in. It is important that how we fund and pay for drugs works for everybody. Change should happen to make sure that both medical services and patients get the treatments and resources they need.
Treatment 10: Aggression, Treatment, Health	<b>[I have to start by saying I utterly disagree with what you've just said. You clearly lack any understanding of how serious this is, and, frankly, you show absolutely no care for the people we represent.]</b> The way we understand diseases and how to treat them has grown over the years. There have been advances in treatments that might help save the lives of people in this country and we should be committed to using them. <b>[But often nothing is done because it is simply deemed not worth the money. I am utterly revolted by the idea that people are not getting treated merely because of some appalling cost-benefit calculation. This is disgusting, deplorable, and inhuman.]</b> It is important that how we fund and pay for drugs works for everybody. <b>[Things must change, and anyone who opposes this cannot claim they care about the well-being of our people.]</b>

Treatment number	Treatment text
Treatment 11: Evidence, Statistics, Health	The way we understand diseases and how to treat them has grown over the years. But in some cases, patients are left without access to the drugs they need because of the costs of treatment. I think some people find it challenging to accept that questions of money should enter decisions about health care, but this is the situation that many hospitals, doctors, and patients find themselves in. <b>[The National Health Service has a fixed budget of only around £110 billion a year in England.]</b> It is important that how we fund and pay for drugs works for everybody. <b>[For example, between 2008 and 2016, one drug increased in price by 12,000%. This sort of increase isn't sustainable, and if the price had only stayed the same, the NHS could have spent £58 million less.]</b> Change should happen to make sure that both medical services and patients get the treatments and resources they need.
Treatment 12: Evidence, Anecdote, Health	The way we understand diseases and how to treat them has grown over the years. <b>[I spoke recently to Eleanor and Michael, a young couple in my constituency, and they told me the story of how Michael was able to get the treatment he needed to be declared cancer free and for his life to be saved.]</b> But in other cases, patients are left without access to the drugs they need because of the costs of treatment. I think some people find it challenging to accept that questions of money should enter decisions about health care, but this is the situation that many hospitals, doctors, and patients find themselves in. It is important that how we fund and pay for drugs works for everybody, <b>[just like it did for Eleanor and Michael].</b> Change should happen to make sure that both medical services and patients get the treatments and resources they need.
Treatment 13: Emotion, Control, Transport	Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. We should aim for all new vehicles made to be run either partially or wholly on electricity. Electric vehicles offer clear benefits for improving local air quality as they produce no exhaust emissions at the street level. The market for electric vehicles is small yet growing. We can be industry leaders in how we produce and use electric vehicles. A variety of different strategies should be employed to encourage their uptake. We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. For instance, by making sure there is sufficient charging infrastructure. If we use electric vehicles, journeys can be greener and safer.
Treatment 14: Emotion, Treatment, Transport	Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. <b>[Imagine stepping outside on a bright, beautiful morning and hearing not engines revving nor choking on polluted air but feeling that simple joy of hearing the birds sing and that rush of fresh air into your lungs. Doesn't this sound amazing? This could be our future, and I so hope that this doesn't have to be a distant dream.]</b> We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer. <b>[We can fill our lives with the simple pleasures of bird song in our ears, fresh air in our lungs, and blue skies ahead.]</b>
Treatment 15: Aggression, Control, Transport	Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. We should aim for all new vehicles made to be run either partially or wholly on electricity. Electric vehicles offer clear benefits for improving local air quality as they produce no exhaust emissions at the street level. The market for electric vehicles is small yet growing. We can be industry leaders in how we produce and use electric vehicles. A variety of different strategies should be employed to encourage their uptake. We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. For instance, by making sure there is sufficient charging infrastructure. If we use electric vehicles, journeys can be greener and safer.

Treatment number	Treatment text
Treatment 16: Aggression, Treatment, Transport	<p><b>[I have to start by saying I utterly disagree with what you've just said. You clearly lack any understanding of how serious this is, and, frankly, you show absolutely no care for the people we represent.]</b> Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. <b>[People are choking to death because we are failing to clean up the toxic air we breathe. It is shameful and nothing has been done. Why?]</b> We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer. <b>[Things must change, and anyone who opposes this cannot claim they care about the well-being of our people.]</b></p>
Treatment 17: Evidence, Statistics, Transport	<p>Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. <b>[Last year, the transport sector accounted for 29.8% of total carbon dioxide emissions.]</b> Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. <b>[In London alone, there are 9,400 premature deaths every year because of poor air quality.]</b> Electric vehicles offer one method of reducing emissions as they produce no air pollution. The market for electric vehicles is small yet growing. <b>[Last year saw the biggest ever annual increase in electric vehicles, with a growth of 175,000 new vehicles, which was up 66% on 2019.]</b> We can be industry leaders in how we produce and use electric vehicles. We should widen accessibility in the use of electric vehicles to make them more practical for those living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer.</p>
Treatment 18: Evidence, Anecdote, Transport	<p>Transport is the largest carbon-emitting sector of the UK economy, and, within this, cars contribute most. Air pollution increases the risk of heart disease, cancer, diabetes, and asthma attacks. Electric vehicles offer one method of reducing emissions as they produce no air pollution. The market for electric vehicles is small yet growing. We can be industry leaders in how we produce and use electric vehicles. <b>[I spoke recently to Eleanor and Michael, a young couple in my constituency who live in a flat in a high-rise building. They told me while they want to make the swap to an electric vehicle, it is just not practical as they do not have easy access to a charging point.]</b> We should widen accessibility in the use of electric vehicles to make them more practical for people, <b>[just like Eleanor and Michael],</b> who live living in urban or built-up areas. If we use electric vehicles, journeys can be greener and safer.</p>

## C.4 Full models

In this section, I present the full regression results from the analysis presented in chapter 5.

### *Unconditional effects*

First, I present the results from the analysis investigating the unconditional effects between style usage and style perceptions, likeability evaluations, and competence evaluations in tables C.4–C.6.

Table C.4: Relationship between style usage and style perceptions

	<i>Dependent variable:</i>		
	Perceived emotion (1)	Perceived aggression (2)	Perceived evidence (3)
Intercept	3.242*** (0.035)	2.918*** (0.041)	3.686*** (0.037)
Emotional Style	0.598*** (0.050)		
Aggressive Style		0.416*** (0.058)	
Anecdotal Style			−0.253*** (0.052)
Observations	1,590	1,587	1,557
R <sup>2</sup>	0.083	0.031	0.015
Adjusted R <sup>2</sup>	0.082	0.031	0.014
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table C.5: Relationship between style usage and likeability evaluations

	<i>Dependent variable:</i>		
	Emotion (1)	Likeability Aggression (2)	Evidence (3)
Intercept	3.430*** (0.030)	3.363*** (0.032)	3.380*** (0.028)
Emotional Style	0.029 (0.042)		
Aggressive Style		−0.271*** (0.045)	
Anecdotal Style			0.125*** (0.039)
Observations	1,487	1,492	1,501
R <sup>2</sup>	0.0003	0.024	0.007
Adjusted R <sup>2</sup>	−0.0004	0.023	0.006
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table C.6: Relationship between style usage and competence evaluations

	<i>Dependent variable:</i>		
	Emotion (1)	Aggression (2)	Evidence (3)
Intercept	3.624*** (0.031)	3.615*** (0.032)	3.641*** (0.031)
Emotional Style	-0.208*** (0.044)		
Aggressive Style		-0.114** (0.046)	
Anecdotal Style			-0.093** (0.043)
Observations	1,538	1,539	1,537
R <sup>2</sup>	0.015	0.004	0.003
Adjusted R <sup>2</sup>	0.014	0.003	0.002
Note:	*p<0.1; **p<0.05; ***p<0.01		

---

*Conditional effects by MP gender*

Second, I present the results from the analysis investigating the conditional effects by MP gender in tables C.7–C.9. The output from model 1 in table C.7 below shows the full regression output for the relationship between MP gender, MP likeability evaluations, and additional control variables as displayed below for the emotion style, and model 2 shows the regression output with the inclusion of additional controls for policy area. The output from model 3 shows the full regression output for the relationship between MP gender, MP likeability evaluations, and additional control variables as displayed below for the aggression style, and model 4 shows the regression output with the inclusion of additional controls for policy area. The output from model 5 shows the full regression output for the relationship between MP gender, MP likeability evaluations, and additional control variables as displayed below for the evidence style, and model 6 shows the regression output with the inclusion of additional controls for policy area. The treatment styles are the emotional style, aggressive style, and anecdotal evidence, the control styles are the non-emotional style, non-aggressive style, and statistical evidence.

In all three models for the likeability outcomes, I have chosen to include respondent gender, age, and degree education as pre-treatment covariates, as, in expectation, they will increase the statistical power of the analysis by better explaining the variation in the likeability outcomes. For instance, a respondent's gender may explain how likeable they deem a politician to be, and, indeed, the results from the conditional analysis by respondent gender presented in tables C.14–C.16 suggest that this is the case. Similarly, a respondent's age may provide a plausible explanation for why they would evaluate a politician delivering, say, an aggressive argument as more or less likeable.

The output from model 1 in table C.8 below shows the full regression output for the relationship between MP gender, MP competence evaluations, and additional control variables as displayed below for the emotion style, and model 2 shows the regression output with the inclusion of additional controls for policy area. The output from model 3 shows the full regression output for the relationship between MP gender, MP competence evaluations, and additional control variables as displayed below for the aggression style, model 4 shows the regression output with the inclusion of additional controls for policy area. The output from model 5 shows the full regression output for the relationship between MP gender, MP competence evaluations, and additional control variables as displayed below for the evidence style, and model 6 shows the regression output with the inclusion of additional controls for policy area. The treatment styles are the emotional style, aggressive style, and anecdotal evidence, the control styles are the non-emotional style, non-aggressive style, and statistical evidence.

Table C.7: Relationship between MP gender and MP likeability evaluations

	Dependent variable:					
	Emotion		Likeability Aggression		Evidence	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	3.651*** (0.085)	3.592*** (0.092)	3.668*** (0.094)	3.525*** (0.099)	3.420*** (0.081)	3.345*** (0.085)
Woman MP	-0.040 (0.059)	-0.038 (0.059)	0.107* (0.063)	0.105* (0.063)	0.082 (0.054)	0.086 (0.054)
Treatment Style	-0.034 (0.057)	-0.033 (0.057)	-0.221*** (0.062)	-0.220*** (0.062)	0.151*** (0.055)	0.148*** (0.055)
Woman Voter	0.188*** (0.042)	0.191*** (0.042)	0.146*** (0.045)	0.145*** (0.045)	0.273*** (0.039)	0.270*** (0.039)
Age	-0.006*** (0.001)	-0.006*** (0.001)	-0.008*** (0.001)	-0.008*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)
Degree Educated	0.018 (0.045)	0.017 (0.045)	-0.085* (0.048)	-0.070 (0.048)	0.046 (0.042)	0.041 (0.041)
Housing Policy		0.077 (0.051)		0.242*** (0.056)		0.136*** (0.045)
Health Policy		0.078 (0.051)		0.164*** (0.053)		0.115** (0.048)
Woman MP*Treatment Style	0.136* (0.082)	0.133 (0.082)	-0.132 (0.088)	-0.133 (0.088)	-0.043 (0.077)	-0.039 (0.076)
Observations	1,487	1,487	1,492	1,492	1,501	1,501
R <sup>2</sup>	0.037	0.039	0.059	0.072	0.058	0.065
Adjusted R <sup>2</sup>	0.033	0.034	0.056	0.067	0.054	0.060

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

As with the likeability outcomes, I have selected pre-treatment covariates that may plausibly explain variation in respondents' competency evaluations. These are respondent gender, age, degree education, and political attention. For instance, political attention has been included as a pre-treatment covariate as respondents with higher levels of political attention may be better equipped to judge the competence of politicians delivering arguments on political issues than respondents with low political attention. Similarly, education has been included as a voter's education level may plausibly explain how competent they deem politicians to be.

The output from model 1 in table C.9 below shows the full regression output for the relationship between MP gender, perceived emotion, and additional control variables as displayed below for the emotion style, and model 2 shows the regression output with the inclusion of additional controls for policy area. The output from model 3 shows the full regression output for the relationship between MP gender, perceived aggression, and additional control variables as displayed below for the aggression style, and model 4 shows the regression output with the inclusion of additional controls for policy area. The output from model 5 shows the full regression output for the relationship between MP gender, perceived evidence, and additional control variables as displayed below for the evidence style, and model 6 shows the regression output with the inclusion of additional controls for policy area. In all

Table C.8: Relationship between MP gender and MP competence evaluations

	Dependent variable:					
	Emotion		Competence Aggression		Evidence	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	3.800*** (0.107)	3.630*** (0.111)	3.669*** (0.114)	3.552*** (0.119)	3.593*** (0.108)	3.477*** (0.110)
Woman MP	-0.040 (0.060)	-0.035 (0.060)	0.020 (0.063)	0.016 (0.063)	0.075 (0.060)	0.083 (0.060)
Treatment Style	-0.267*** (0.060)	-0.263*** (0.059)	-0.111* (0.063)	-0.112* (0.063)	-0.052 (0.061)	-0.059 (0.061)
Woman Voter	0.295*** (0.044)	0.301*** (0.044)	0.217*** (0.047)	0.218*** (0.047)	0.278*** (0.044)	0.271*** (0.044)
Age	-0.008*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)
Degree Educated	-0.036 (0.048)	-0.038 (0.047)	-0.024 (0.050)	-0.010 (0.050)	0.054 (0.048)	0.044 (0.047)
Political Attention	0.011 (0.010)	0.010 (0.010)	0.026** (0.010)	0.026** (0.010)	0.020** (0.010)	0.020** (0.010)
Housing Policy		0.191*** (0.053)		0.192*** (0.057)		0.240*** (0.050)
Health Policy		0.274*** (0.053)		0.142*** (0.054)		0.152*** (0.053)
Woman MP*Treatment Style	0.124 (0.085)	0.113 (0.084)	-0.039 (0.090)	-0.035 (0.090)	-0.077 (0.085)	-0.068 (0.084)
Observations	1,538	1,538	1,539	1,539	1,537	1,537
R <sup>2</sup>	0.070	0.087	0.038	0.046	0.048	0.063
Adjusted R <sup>2</sup>	0.066	0.081	0.034	0.040	0.044	0.058

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

three models, the treatment styles are the emotional style, aggressive style, and anecdotal evidence. The control styles are the non-emotional style, the non-aggressive style, and statistical evidence.

Finally, as with the above models, I have again included pre-treatment covariates that may plausibly explain respondents' style perceptions. For emotion and aggression, these are respondent gender, left-right placement, and age. For evidence, this is respondent gender, age, and degree education. Left-right placement has been included in the models for perceived emotion and aggression as work has suggested that left-wing ideology has historically been associated with greater emotionality than right-wing, and therefore voters' left-right positioning may explain variation in the emotion and aggression perceptions (e.g., [Salmela and Von Scheve, 2018](#)). Similarly, I include degree education in the perceived evidence model, as a respondent's degree education may plausibly explain how they evaluate the extent to which they perceive anecdotal or statistical arguments to be evidence-based.



Table C.9: Relationship between MP gender and style perceptions

	<i>Dependent variable:</i>					
	Perceived emotion		Perceived aggression		Perceived evidence	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	3.144*** (0.125)	2.905*** (0.135)	2.768*** (0.150)	2.714*** (0.157)	3.606*** (0.111)	3.468*** (0.115)
Woman MP	-0.107 (0.081)	-0.099 (0.081)	0.088 (0.095)	0.082 (0.095)	0.041 (0.074)	0.049 (0.073)
Treatment Style	0.679*** (0.081)	0.692*** (0.080)	0.567*** (0.094)	0.565*** (0.094)	-0.206*** (0.075)	-0.213*** (0.075)
Woman Voter	0.065 (0.058)	0.070 (0.058)	-0.091 (0.068)	-0.089 (0.068)	0.220*** (0.053)	0.212*** (0.053)
Left-Right Placement	0.013 (0.021)	0.022 (0.021)	-0.012 (0.025)	-0.012 (0.025)		
Age	0.00004 (0.002)	0.0002 (0.002)	0.003 (0.002)	0.003 (0.002)	-0.001 (0.002)	-0.001 (0.002)
Housing Policy		0.187*** (0.072)		0.059 (0.084)		0.278*** (0.061)
Health Policy		0.352*** (0.072)		0.110 (0.080)		0.185*** (0.065)
Degree Educated					-0.045 (0.056)	-0.056 (0.056)
Woman MP*Treatment Style	0.067 (0.116)	0.048 (0.115)	-0.174 (0.135)	-0.168 (0.135)	-0.081 (0.104)	-0.071 (0.103)
Observations	1,226	1,226	1,220	1,220	1,557	1,557
R <sup>2</sup>	0.112	0.129	0.044	0.046	0.028	0.041
Adjusted R <sup>2</sup>	0.107	0.123	0.040	0.040	0.024	0.036

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## C.5 Multiple comparisons correction

In the empirical strategy carried out in chapter 5, there are numerous statistical tests are conducted, and there is risk of the multiple comparisons problem. To quote [Gelman, Hill and Yajima \(2012, 189-190\)](#), the multiple comparisons problem “is that the probability of a researcher wrongly concludes that there is at least one statistically significant effect across a set of tests, even when in fact there is nothing going on, increases with each additional test”. In other words, when conducting a very large number of tests any one might be “significant” by chance alone, meaning that the  $p$ -values are unlikely to capture the true Type 1 error rate. Therefore, the probability of falsely rejecting a null hypothesis which is correct increases with each additional test.

To assess whether the results I present in the main body are robust, I here carry out subsequent analyses with adjusted  $p$ -values which control for the False Discovery Rate using the Benjamini-Hochberg procedure ([Benjamini and Hochberg, 1995](#)). An alternative common approach is the Bonferroni correction. Under this approach, the  $p$ -value at which a test is evaluated is based on the total number of tests performed. Practically speaking, the  $p$ -value is calculated as the original  $p$ -value divided by the number of tests performed. However, the Bonferroni correction assumes independence between tests conducted which is clearly inappropriate in the case of this experimental design, where any form of gender-bias driving perceptions of women’s use of, say, aggression is also likely to inform how respondents perceive and evaluate women’s use of, say, statistical evidence. Further, by targeting the Type 1 error problem, the Bonferroni correction increases the likelihood of Type 2 errors. By changing the  $p$ -value needed to reject the null hypothesis, it increases the number of instances where the null is not rejected when it is in fact false and should have been. As [Gelman, Hill and Yajima \(2012, 192\)](#) argue: “the Bonferroni correction can severely reduce our power to detect an important effect”.

I instead opt to use the Benjamini-Hochberg ([Benjamini and Hochberg, 1995](#)) procedure which is less stringent than the Bonferroni correction, and is more appropriate in the context of this experimental design. In practice, when using the Benjamini-Hochberg procedure, an  $\alpha$  level is selected, which here is 0.05. Next, the  $p$ -values of all hypotheses tests are ordered. Then I identify the largest  $p$ -value which satisfies the criteria  $p_k \leq \frac{k}{m}\alpha$  where  $k$  is the  $p$ -value’s index. This test, and all tests with smaller  $p$ -values are declared significant. I apply this procedure to all of the unconditional and conditional models run in the main body of the text. The relevant  $p$ -values in the unconditional models are those that relate to the differences between treatment and control styles –  $\beta_1$ . The relevant  $p$ -values in the conditional models are those that relate to the differences in the effect between treatment and control

styles for men and women MPs –  $\beta_3$ .

Table C.10 shows the results from this procedure where I report the unadjusted and Benjamini–Hochberg adjusted  $p$ -values for the coefficients described above, whether they are significant or not, and whether the correction changes the significance. As is clear from the table, the results presented in chapter 5 are robust to the correction in all instances except for the model showing the unconditional relationship between MP’s use of evidence and competence evaluations (unconditional evidence competence). After the multiple comparisons correction is applied, this coefficient is no longer statistically significant.

Table C.10: Comparison between unadjusted and Benjamini-Hochberg adjusted p-values

Model name	Unadjusted p-value	Significant?	Adjusted p-value	Significant?	Difference?
Unconditional emotion perception	0.00	Yes	0.00	Yes	No
Unconditional aggression perception	0.00	Yes	0.00	Yes	No
Unconditional evidence perception	0.00	Yes	0.00	Yes	No
Unconditional emotion likeability	0.49	No	0.59	No	No
Unconditional aggression likeability	0.00	Yes	0.00	Yes	No
Unconditional evidence likeability	0.00	Yes	0.00	Yes	No
Unconditional emotion competence	0.00	Yes	0.00	Yes	No
Unconditional aggression competence	0.01	Yes	0.03	Yes	No
Unconditional evidence competence	0.03	Yes	0.07	No	Yes
Conditional emotion likeability	0.10	No	0.20	No	No
Conditional aggression likeability	0.14	No	0.24	No	No
Conditional evidence likeability	0.58	No	0.61	No	No
Conditional emotion competence	0.15	No	0.24	No	No
Conditional aggression competence	0.67	No	0.67	No	No
Conditional evidence competence	0.37	No	0.51	No	No
Conditional emotion perception	0.56	No	0.61	No	No
Conditional aggression perception	0.20	No	0.29	No	No
Conditional evidence perception	0.44	No	0.56	No	No

## C.6 Pooled models

In this section, I present the analysis of two sets of analysis where I pool the styles together. To do so, I first create a variable for whether a style is **female stereotype-congruent**, which takes the value of 0 for female stereotype-incongruent styles (non-emotional style, aggressive style, and statistical evidence) and 1 for female stereotype-congruent styles (emotional style, non-aggressive style, and anecdotal evidence). To assess how politicians using styles which are congruent with female stereotypes affecting voters' likeability and competence evaluations of them. To do so for the likeability outcomes, I pool all styles together and estimate the following:

$$Likeability_{i(j)} = \alpha + \beta_1 FemaleStereotypeCongruent_j + \epsilon_i \quad (C.1)$$

where  $\alpha$  describes the average likeability evaluations in female stereotype-incongruent styles (non-emotional style, aggressive style, and statistical evidence), and  $\alpha + \beta_1$  describes the same quantities for female stereotype-congruent styles (emotional style, non-aggressive style, and anecdotal evidence).

I can, of course, identify the same quantity for the competence outcomes. I therefore also estimate:

$$Competence_{i(j)} = \alpha + \beta_1 FemaleStereotypeCongruent_j + \epsilon_i \quad (C.2)$$

where  $\alpha$  describes the average competence evaluations in female stereotype-incongruent styles (non-emotional style, aggressive style, and statistical evidence), and  $\alpha + \beta_1$  describes the same quantities for female stereotype-congruent styles (emotional style, non-aggressive style, and anecdotal evidence). As there are multiple observations per respondent, I cluster standard errors in both models at the respondent level. The results are presented in table C.11.

Table C.11: Relationship between female stereotype congruent styles, likeability, and competence evaluations

	Dependent variable:	
	Likeability (1)	Competence (2)
Intercept	3.299*** (0.017)	3.589*** (0.018)
Female Stereotype Congruent	0.144*** (0.024)	-0.062** (0.026)
Observations	4,480	4,614
R <sup>2</sup>	0.008	0.001
Adjusted R <sup>2</sup>	0.007	0.001
Note:	*p<0.1; **p<0.05; ***p<0.01	

In the left-hand column, I show the results for the likeability outcomes. Here, the coefficient for female-stereotype-congruent (emotional style, non-aggressive style, and anecdotal evidence) is positive and significant, suggesting that politicians are evaluated as *more* likeable when they express styles which are congruent with “communal” stereotypes associated with women. This finding perhaps seems intuitive, given that communal stereotypes are associated with being warm, kind, emotional, and people-oriented (Eagly and Wood, 2012; Schneider and Bos, 2019).

In the right-hand column, I show the results for the competence outcomes. Here, the coefficient for female-stereotype-congruent (emotional style, non-aggressive style, and anecdotal evidence) is negative and significant, suggesting that politicians are evaluated as *less* competent when they express styles which are congruent with “communal” stereotypes associated with women. That voters find politicians to be more competent when they express styles consistent with agentic stereotypes again seems to be an intuitive finding given the compatibility between agentic stereotypes and leadership stereotypes (Bauer, 2017). The takeaway from these findings, therefore, is that style usage represents a trade-off for politicians. While politicians gain in the likeability assessments when they use “communal” styles, they lose out in their competence evaluations.

Next, in the main analysis conditional analysis by MP gender presented in chapter 5, I analyse each style separately. I here present the results from an analysis where I pool all styles together and compare each treatment back to the control arguments. To do so, I first create a categorical variable for the **style type** of an argument, which takes the values of control (0), emotion (1), aggression (2), statistics (3), and anecdote (4). With this variable in hand, I estimate the following:

$$\begin{aligned}
 Likeability_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 Emotion_j + \beta_3 Aggression_j + \\
 & \beta_4 Statistics_j + \beta_5 Anecdote_j + \beta_6 (WomanMP \cdot Emotion) + \\
 & \beta_7 (WomanMP \cdot Aggression) + \beta_8 (WomanMP \cdot Statistics) + \\
 & \beta_9 (WomanMP \cdot Anecdote) + \gamma X_{i(j)} + \epsilon_i
 \end{aligned}
 \tag{C.3}$$

where  $\beta_1$  describes the difference in likeability evaluations between women and men MPs in the control condition.  $\beta_2$ - $\beta_5$  describe the effect of each style type on MP likeability evaluations compared to the control condition  $\alpha$ .  $\beta_6$ - $\beta_9$  describe the difference in the effect of each style type on likeability evaluations for women MPs compared to men. My argument is that women MPs will be rewarded for expressing styles that are congruent with female stereotypes, while they will be punished for express-

ing female stereotype-incongruent styles. As such, I expect that the coefficients for  $\beta_6$  and  $\beta_9$  to be positive, and the coefficients for  $\beta_7$  and  $\beta_8$  to be negative.  $X_i$  represents a control for each policy area (transport, housing, and health). As there are multiple observations per respondent, I cluster standard errors at the respondent level  $i$ .

I also estimate the same model for the competence outcome:

$$\begin{aligned} Competence_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 Emotion_j + \beta_3 Aggression_j + \\ & \beta_4 Statistics_j + \beta_5 Anecdote_j + \beta_6 (WomanMP \cdot Emotion) + \\ & \beta_7 (WomanMP \cdot Aggression) + \beta_8 (WomanMP \cdot Statistics) + \\ & \beta_9 (WomanMP \cdot Anecdote) + \gamma X_{i(j)} + \epsilon_i \end{aligned} \quad (C.4)$$

where  $\beta_1$ - $\beta_9$  represent the same quantities as above. Again my argument is that women MPs will be rewarded for stereotype-congruent behaviour and punished for stereotypes-incongruent behaviour. As such, I expect that  $\beta_7$  will be negative, and  $\beta_6$ ,  $\beta_8$ , and  $\beta_9$  will be positive.  $X_i$  again represents a control for each policy area (transport, housing, and health). I cluster standard errors at the respondent level  $i$ . The results from both models are presented in table C.12.

Looking first at the left-hand column of table C.12, which shows the results for the likeability outcome, I see that the coefficient for  $\beta_1$  (Woman MP) is insignificant. This suggests there is no difference in likeability evaluations between men and women MPs in the control condition.  $\beta_2$  (Emotion) and  $\beta_4$  (Statistics) are also both insignificant, suggesting that among men, the use of emotional and statistical styles have no statistically significant effect on voters' evaluations of men's likeability relative to the control. However,  $\beta_3$  (Aggression) is negative and significant, suggesting that when men deliver aggressive arguments they are perceived as less likeable. Similarly,  $\beta_5$  (Anecdote) is positive and significant, and voters therefore ascribe men higher likeability evaluations when using anecdotes compared to the control. The table also shows positive and significant coefficients for both housing and health policy areas, which suggests that voters find politicians to be more likeable in these issue areas compared to in transport, the baseline.

Interestingly, the coefficients for each of the interaction terms –  $\beta_6$ - $\beta_9$  in equation 5.5 described above – are all statistically insignificant. The results from this analysis therefore support the analysis in the main body of the chapter: that although style usage influences voters' evaluations of politicians' likeability, these evaluations are *not* gendered.

Turning to the right-hand column of table C.12, I present the results from equation C.4.  $\beta_1$  (Woman

Table C.12: Relationship between MP gender, style usage, likeability, and competence evaluations

	<i>Dependent variable:</i>	
	Likeability	Competence
	(1)	(2)
Intercept	3.287*** (0.034)	3.492*** (0.036)
Woman MP	0.028 (0.042)	-0.019 (0.044)
Emotion	0.025 (0.051)	-0.253*** (0.054)
Aggression	-0.265*** (0.050)	-0.106** (0.053)
Statistics	-0.039 (0.052)	-0.016 (0.054)
Anecdote	0.110** (0.051)	-0.075 (0.054)
Housing Policy	0.156*** (0.030)	0.206*** (0.031)
Health Policy	0.123*** (0.030)	0.190*** (0.031)
Woman MP*Emotion	0.068 (0.072)	0.093 (0.076)
Woman MP*Aggression	-0.077 (0.073)	-0.020 (0.076)
Woman MP*Statistics	0.057 (0.073)	0.096 (0.076)
Woman MP*Anecdote	0.009 (0.073)	0.026 (0.076)
Observations	4,480	4,614
R <sup>2</sup>	0.034	0.021
Adjusted R <sup>2</sup>	0.032	0.018
Note:	*p<0.1; **p<0.05; ***p<0.01	



MP) is again insignificant, and therefore also appears to be no difference in competence evaluations between men and women MPs in the control condition.  $\beta_2$  (Emotion) and  $\beta_3$  (Aggression) are, however, both negative and significant, and voters therefore evaluate men as less competent when they use emotional arguments or aggressive arguments compared to the control. Both  $\beta_4$  (Statistics) and  $\beta_5$  (Anecdote) are insignificant. As in the likeability model, the coefficients for housing and health are both positive and significant, suggesting that voters also ascribe higher competence evaluations to politicians delivering arguments on these issue areas compared to transport.

The coefficients for the each of the interaction terms –  $\beta_6$ – $\beta_9$  in equation C.4 – are again statistically insignificant. There is no evidence that women MPs in particular are punished in competence evaluations when they violate stereotypes, nor are they rewarded when they conform to stereotypes, compared to men. Overall, the results from both models support the findings presented in chapter 5.

## C.7 MP gender and policy area

While not the main quantity of interest, it is possible that men and women MPs may receive differential evaluations in likeability and competence depending on the policy area in question. A rich body of literature has shown that women are stereotyped to be better suited to and more qualified on issues that relate to feminine communal stereotypes, while men are expected to instead have more authority on issues related to the masculine, agentic, and assertive stereotypes to which they are associated (Huddy and Terkildsen, 1993a; Kahn, 1996; Lawless, 2004; McDermott, 1997; Schneider and Bos, 2019). Other work has shown that women candidates are more successful when they run on “women’s issues” (Ennsner-Jedenastik, 2017; Herrnson, Lay and Stokes, 2003). Once in office, women politicians introduce and advocate for legislation on “feminine” social policy issues (Schwindt-Bayer, 2006; Swers, 2002; Thomas, 1994), disproportionately participate in debates on women’s issues (Bäck and Debus, 2019; Catalano, 2009), and may raise traditional women’s issues in parliamentary speeches more than men (Bailer et al., 2022). Further, past experimental work has shown, for instance, that the policy context may have important implications for the power of gender stereotypes (Anzia and Bernhard, 2022; Holman, Merolla and Zechmeister, 2011, 2016, 2017). Further, women have been said to be more persuasive on “feminine” policy areas and men on “masculine” policy areas, although recent experimental work finds little evidence to support this (Anderson-Nilsson and Clayton, 2021; Searles et al., 2020).

In the experimental design described in chapter 5, I vary the policy area of the arguments, focusing on housing, health, and transport. While there is some debate in the literature on whether housing and transport policies are stereotypically “feminine” and “masculine” (Krook and O’Brien, 2012), health is an area that has commonly been associated with “feminine” stereotypes of being communal and caring (Catalano, 2009; Kittilson, 2011; Norris, 1996). As such, it may be the case that women are evaluated as more likeable and competent than men on health as opposed to transport or housing policy issues. To assess whether this is the case, I estimate the following:

$$\begin{aligned}
 \textit{Likeability}_{i(j)} = & \alpha + \beta_1 \textit{WomanMP}_j + \beta_2 \textit{Housing}_j + \beta_3 \textit{Health}_j + \\
 & \beta_4 (\textit{WomanMP} \cdot \textit{Housing}) + \\
 & \beta_5 (\textit{WomanMP} \cdot \textit{Health}) + \\
 & \gamma \textit{X}_{i(j)} + \epsilon_i
 \end{aligned}
 \tag{C.5}$$

where  $\beta_1$  describes the difference in likeability between women and men MPs in the transport condition.  $\beta_2$ - $\beta_3$  describe the effect of housing and health policies on MP likeability evaluations compared to transport among men.  $\beta_4$ - $\beta_4$  describe the difference in the effect of housing and health policies compared to transport on likeability for women MPs compared to men.  $X_i$  represents a control for each style (emotion, aggression, and evidence). As there are multiple observations per respondent, I cluster standard errors at the respondent level  $i$ .

Similarly, I can estimate the same quantity for the competence outcomes:

$$\begin{aligned}
 Competence_{i(j)} = & \alpha + \beta_1 WomanMP_j + \beta_2 Housing_j + \beta_3 Health_j + \\
 & \beta_4 (WomanMP \cdot Housing) + \\
 & \beta_5 (WomanMP \cdot Health) + \\
 & \gamma X_{i(j)} + \epsilon_i
 \end{aligned}
 \tag{C.6}$$

where  $\beta_1$ - $\beta_6$  describe the same quantities described above. I present the results from both models in table C.13. The coefficients that enable me to see whether voters are awarding differentially likeability and competence evaluations for men and women MPs are the coefficient for  $\beta_1$  (Woman MP),  $\beta_4$  (Woman MP\*Housing), and  $\beta_5$  (Woman MP\*Health). As is clear from the table, all three coefficients are statistically insignificant, and, consequently, I see no evidence of differentially evaluations for men and women according to policy area.

Table C.13: Relationship between MP gender, policy area, likeability, and competence evaluations

	<i>Dependent variable:</i>	
	Likeability (1)	Competence (2)
Intercept	3.339*** (0.033)	3.387*** (0.037)
Woman MP	0.017 (0.044)	-0.012 (0.047)
Housing Policy	0.153*** (0.038)	0.188*** (0.042)
Health Policy	0.090** (0.040)	0.171*** (0.042)
Aggression	-0.215*** (0.025)	0.045* (0.027)
Evidence	0.007 (0.021)	0.089*** (0.026)
Woman MP*Housing Policy	0.003 (0.063)	0.037 (0.067)
Woman MP*Health Policy	0.067 (0.063)	0.036 (0.065)
Observations	4,480	4,614
R <sup>2</sup>	0.023	0.013
Adjusted R <sup>2</sup>	0.022	0.011
Note:	*p<0.1; **p<0.05; ***p<0.01	

## C.8 Heterogeneous effects by voter gender

How might (gendered) perceptions and evaluations of the styles politicians use vary by voter gender? In the pre-analysis plan (Hargrave, 2021), I stated I would carry out an exploratory analysis into how the treatment effects may differ by respondent characteristics. In this section, I carry out two sets of analysis to investigate how the effects may vary by voter gender. First, to assess whether men and women voters are equivalently sensitive to the styles that politicians use. Second, to assess whether men and women voters are differentially sensitive to the extent to which women politicians conform to stereotype-congruent behaviours. I carry out each below.

### *Conditional relationships by voter gender*

Here I am interested in identifying whether men and women voters' evaluations and perceptions are equivalently sensitive to the styles politicians use. That is, might certain kinds of voters' evaluations of the likeability or competence of politicians be more affected by the styles used? I assess whether this is the case by interacting voter gender and style prevalence for each of the outcomes. Tables C.14–C.16 show the full results.

As described in the chapter 5, there are few overall differences in how politicians' style usage influences the perceptions and evaluations of men and women voters. For emotion (table C.14) none of the interaction terms are significant, as is also the case for aggression (table C.15). However, for evidence (table C.16) there are some differences. While men do not find politicians' use of anecdotal evidence as less competent or evidence-based than statistical evidence, women voters do. Further, while the use of anecdotal evidence improves men voters' likeability assessments of politicians relative to the use of statistical evidence, this is not the case for women voters. Therefore, to the extent that there are differences in how men and women voters evaluate politicians' use of styles, the differences are concentrated among the evidence style.

Table C.14: Relationship between voter gender and likeability evaluations, competence evaluations, and emotion perceptions

	<i>Dependent variable:</i>		
	Likeability (1)	Competence (2)	Perceived Emotion (3)
Intercept	3.282*** (0.046)	3.409*** (0.047)	3.167*** (0.055)
Woman Voter	0.255*** (0.060)	0.373*** (0.062)	0.130* (0.072)
Emotional Style	0.074 (0.064)	-0.148** (0.066)	0.624*** (0.077)
Woman Voter*Emotional Style	-0.079 (0.084)	-0.107 (0.087)	-0.046 (0.101)
Observations	1,487	1,538	1,590
R <sup>2</sup>	0.018	0.049	0.086
Adjusted R <sup>2</sup>	0.016	0.047	0.084

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table C.15: Relationship between voter gender and likeability evaluations, competence evaluations, and aggression perceptions

	<i>Dependent variable:</i>		
	Likeability (1)	Competence (2)	Perceived Aggression (3)
Intercept	3.256*** (0.049)	3.497*** (0.049)	2.939*** (0.063)
Woman Voter	0.188*** (0.064)	0.207*** (0.065)	-0.037 (0.083)
Aggressive Style	-0.276*** (0.069)	-0.144** (0.070)	0.428*** (0.090)
Woman Voter*Aggressive Style	0.002 (0.091)	0.041 (0.092)	-0.018 (0.118)
Observations	1,492	1,539	1,587
R <sup>2</sup>	0.035	0.020	0.032
Adjusted R <sup>2</sup>	0.033	0.018	0.030

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table C.16: Relationship between voter gender and likeability evaluations, competence evaluations, and evidence perceptions

	<i>Dependent variable:</i>		
	Likeability (1)	Competence (2)	Perceived Evidence (3)
Intercept	3.167*** (0.042)	3.416*** (0.047)	3.489*** (0.057)
Woman Voter	0.367*** (0.055)	0.382*** (0.061)	0.335*** (0.075)
Anecdote	0.215*** (0.059)	0.028 (0.066)	-0.126 (0.080)
Woman Voter*Anecdote	-0.151* (0.078)	-0.200** (0.087)	-0.212** (0.105)
Observations	1,501	1,537	1,557
R <sup>2</sup>	0.045	0.033	0.029
Adjusted R <sup>2</sup>	0.043	0.031	0.027

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

*Conditional relationships by voter and MP gender*

As described in chapter 5, I am also interested in identifying whether men and women voters are differentially sensitive to the extent to which politicians conform to or violate stereotype-congruent behaviours. Here, I assess whether this is the case by subsetting the data into men and women voters and replicating the analysis from chapter 5. Tables C.17–C.19 present the results from each of the outcomes arranged by the styles.

While for aggression (table C.18) and evidence (table C.19) there do not seem to be any differences between men and women voters, this is not the case for emotion (table C.17). For likeability, among women voters, I see that the coefficient on the interaction term is positive and significant, suggesting that women politicians *in particular* are rewarded for expressing emotional styles instead of non-emotional styles. I do not find the same effect among men voters. For competence, the interaction term is again positive and significant, suggesting that while women voters find emotional politicians to be less competent than non-emotional politicians, they give women MPs *more* of a competency reward than men MPs. I again see no such effect among men voters. Finally, for perceived styles I again see that the interaction term is positive and significant. While women voters perceive both men and women politicians as more emotional when they use emotional styles than non-emotional styles, they perceive women MPs *in particular* to be more emotional than men MPs. I, again, do not find the same effect among men voters.

Therefore, women voters give a bigger likeability and competence reward to women politicians who are emotional and perceive women MPs as more emotional than men MPs. I find no evidence of these effects among men respondents. Therefore, to the extent that there is any evidence of differential evaluations between men and women voters of men and women politicians for the styles they use, it seems that this effect is concentrated amongst women voters for the emotion style.

Table C.17: Relationship between MP gender, voter gender, and likability, competence, and perceived emotion for the emotion style

	<i>Dependent variable:</i>					
	Likeability		Competence		Perceived emotion	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Intercept	3.284*** (0.067)	3.571*** (0.050)	3.394*** (0.072)	3.832*** (0.050)	3.103*** (0.079)	3.378*** (0.061)
Woman MP	-0.004 (0.098)	-0.073 (0.074)	0.032 (0.104)	-0.105 (0.073)	0.134 (0.115)	-0.170* (0.089)
Emotional Style	0.062 (0.096)	-0.118* (0.071)	-0.088 (0.104)	-0.406*** (0.071)	0.775*** (0.114)	0.442*** (0.087)
Woman MP*Emotional Style	0.023 (0.137)	0.231** (0.103)	-0.119 (0.148)	0.310*** (0.102)	-0.304* (0.162)	0.285** (0.125)
Observations	620	867	647	891	665	925
R <sup>2</sup>	0.002	0.007	0.007	0.038	0.087	0.090
Adjusted R <sup>2</sup>	-0.003	0.003	0.003	0.035	0.083	0.087

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table C.18: Relationship between MP gender, voter gender, and likability, competence, and perceived aggression for the aggression style

	<i>Dependent variable:</i>					
	Likeability		Competence		Perceived aggression	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Intercept	3.236*** (0.072)	3.365*** (0.058)	3.503*** (0.077)	3.685*** (0.054)	2.854*** (0.092)	2.899*** (0.075)
Woman MP	0.041 (0.101)	0.156* (0.082)	-0.012 (0.109)	0.038 (0.077)	0.170 (0.129)	0.006 (0.106)
Aggressive Style	-0.177* (0.102)	-0.215*** (0.079)	-0.074 (0.111)	-0.109 (0.075)	0.544*** (0.131)	0.494*** (0.103)
Woman MP*Aggressive Style	-0.195 (0.143)	-0.113 (0.114)	-0.139 (0.157)	0.017 (0.108)	-0.232 (0.185)	-0.181 (0.149)
Observations	627	865	641	898	667	920
R <sup>2</sup>	0.027	0.030	0.008	0.005	0.034	0.035
Adjusted R <sup>2</sup>	0.022	0.027	0.003	0.002	0.029	0.032

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

Table C.19: Relationship between MP gender, voter gender, and likability, competence, and perceived evidence for the evidence style

	<i>Dependent variable:</i>					
	Likeability		Competence		Perceived evidence	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Intercept	3.116*** (0.066)	3.491*** (0.048)	3.362*** (0.078)	3.764*** (0.050)	3.403*** (0.090)	3.843*** (0.065)
Woman MP	0.096 (0.090)	0.085 (0.068)	0.102 (0.107)	0.067 (0.070)	0.161 (0.123)	-0.036 (0.091)
Anecdote	0.222** (0.093)	0.098 (0.067)	0.084 (0.110)	-0.144** (0.071)	-0.076 (0.127)	-0.293*** (0.091)
Woman MP*Anecdote	-0.016 (0.127)	-0.066 (0.096)	-0.105 (0.150)	-0.057 (0.100)	-0.096 (0.172)	-0.095 (0.129)
Observations	634	867	648	889	657	900
R <sup>2</sup>	0.021	0.004	0.002	0.014	0.006	0.032
Adjusted R <sup>2</sup>	0.016	0.0004	-0.003	0.011	0.002	0.029

Note:

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01



## C.9 Power analysis

Figure C.5 shows the results of a power analysis for the main effects outlined in chapter 5. The power analysis was conducted after the experiment and, as such, the effect size remains fixed. To construct the power analysis, I simulated the data collection process for the fixed sample of 1,600 respondents I have available for different hypothetical standardised effect sizes for the interaction between style treatment and MP gender (from very small – 0.01 standard deviations – up to large – 0.8 – standard deviations according to conventional Cohen's  $d$  standards). Note that here, to simplify the power analysis, I treat the styles as separate factorial designs, where I have a treatment and control condition and a binary moderator (MP gender) and I treat the outcome as a continuous variable, as opposed to a 5-point Likert scale.

On the basis of an analysis with 1,600 respondents – the fixed number I have available – the power analysis suggests that I am well powered to detect a standardised interaction effect size of 0.28 with 80% power. While this is reasonably small by conventional Cohen's  $d$  standards, it is also roughly comparable to the effects I estimate for the style treatments for men MPs across the various styles and outcomes. This suggests that I cannot confidently rule out the possibility of non-negligible interaction effects, but any interactions that do exist are nevertheless unlikely to be large relative to the overall variance in the outcome variables I study.

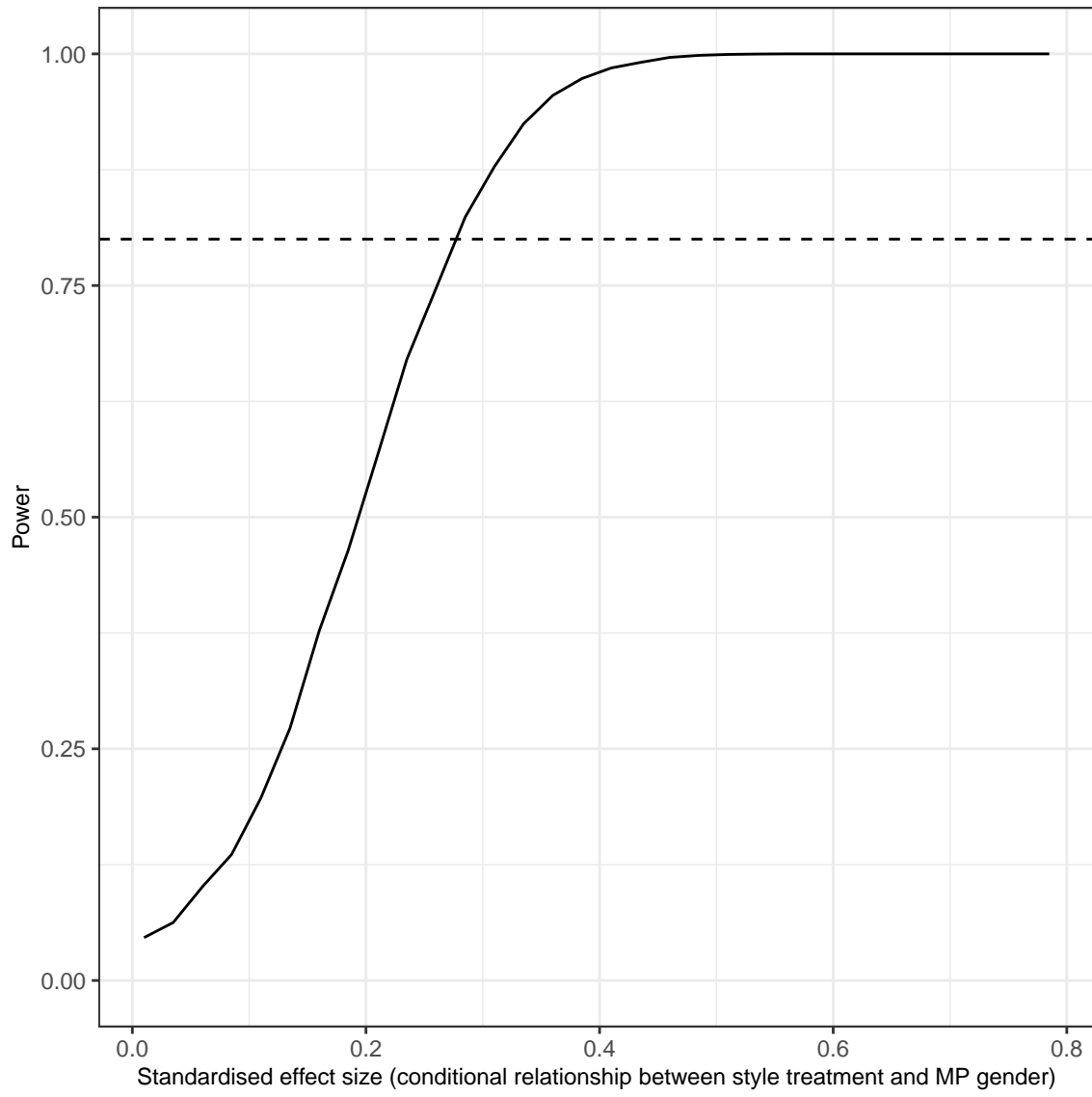


Figure C.5: Power analysis

# Bibliography

- Adams, James F., Samuel Merrill III and Bernard Grofman. 2005. *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. Cambridge, UK: Cambridge University Press.
- Aday, Sean and James Devitt. 2001. *Style over substance. Newspaper coverage of female candidates: Spotlight on Elizabeth Dole*. Washington, DC: The Women's Leadership Fund.
- Alexander, Deborah and Kristi Andersen. 1993. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46(3):527–545.
- Allen, Peter. 2013. "Linking Pre-Parliamentary Political Experience and the Career Trajectories of the 1997 General Election Cohort." *Parliamentary Affairs* 66:685–707.
- Allen, Peter, David Cutts and Madelaine Winn. 2016. "Understanding Legislator Experiences of Family-Friendly Working Practices in Political Institutions." *Politics and Gender* 12(3):549–572.
- Anderson-Nilsson, Georgia and Amanda Clayton. 2021. "Gender and Policy Persuasion." *Political Science Research and Methods* 9(4):818–831.
- Annesley, Claire, Isabelle Engeli and Francesca Gains. 2015. "The Profile of Gender Equality Issue Attention in Western Europe." *European Journal of Political Research* 54(3):525–542.
- Anzia, Sarah F. and Christopher R. Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55(3):478–493.
- Anzia, Sarah F. and Rachel Bernhard. 2022. "Gender Stereotyping and the Electoral Success of Women Candidates: New Evidence from Local Elections in the United States." *British Journal of Political Science* (Forthcoming):1–20.
- Atkins, Judi and Alan Finlayson. 2013. "'... A 40-Year-Old Black Man Made the Point to Me': Everyday

- Knowledge and the Performance of Leadership in Contemporary British Politics." *Political Studies* 61(1):161–177.
- Bäck, Hanna and Marc Debus. 2019. "When Do Women Speak? A Comparative Analysis of the Role of Gender in Legislative Debates." *Political Studies* 67(3):576–596.
- Bäck, Hanna, Marc Debus and Jochen Müller. 2014. "Who Takes the Parliamentary Floor? The Role of Gender in Speech-Making in the Swedish Riksdag." *Political Research Quarterly* 67(3):504–518.
- Bäck, Hanna, Marc Debus and Jorge M. Fernandes. 2021. *The Politics of Legislative Debates*. Oxford: Oxford University Press.
- Bailer, Stefanie, Christian Breunig, Nathalie Giger and Andreas Wüst. 2022. "The Diminishing Value of Minority Representation: Between Group Representation and Individual Career Paths." *British Journal of Political Science* 52(2):535–552.
- Bailer, Stefanie and Tamaki Ohmura. 2018. "Exploring, Maintaining, and Disengaging—The Three Phases of a Legislator's Life." *Legislative Studies Quarterly* 43(3):493–520.
- Bakan, David. 1966. *The Duality of Human Existence*. Chicago, US: Rand McNally.
- Ban, Pamela, Justin Grimmer, Jaclyn Kaslovsky and Emily West. 2022. "How Does the Rising Number of Women in the U.S. Congress Change Deliberation? Evidence from House Committee Hearings." *Quarterly Journal of Political Science* 17(3).
- Banaji, Mahzarin R. and Anthony G. Greenwald. 1995. "Implicit Gender Stereotyping in Judgments of Fame." *Journal of Personality and Social Psychology* 68(2):181–198.
- Banaji, Mahzarin R. and Curtis D. Hardin. 1996. "Automatic Stereotyping." *Psychological Science* 7(3):136–141.
- Banaji, Mahzarin R., Curtis Hardin and Alexander J. Rothman. 1993. "Implicit Stereotyping in Person Judgment." *Journal of Personality and Social Psychology* 65(2):272–281.
- Banwart, Mary Christine, Dianne G. Bystrom and Terry Robertson. 2003. "From the Primary to the General Election: A Comparative Analysis of Candidate Media Coverage in Mixed-Gender 2000 Races for Governor and U.S. Senate." *American Behavioral Scientist* 46(5):658–676.
- Banwart, Mary Christine and Mitchell S. McKinney. 2005. "A Gendered Influence in Campaign Debates?"

- Analysis of Mixed-gender United States Senate and Gubernatorial Debates." *Communication Studies* 56(4):353–373.
- Barnes, Tiffany. 2016. *Gendering Legislative Behaviour: Institutional Constraints and Collaboration*. Cambridge: Cambridge University Press.
- Barnes, Tiffany D. and Diana Z. O'Brien. 2018. "Defending the Realm: The Appointment of Female Defense Ministers Worldwide." *American Journal of Political Science* 62(2):355–368.
- Barnes, Tiffany D. and Emily Beaulieu. 2019. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative Political Studies* 52(1):134–167.
- Bateson, Regina. 2020. "Strategic Discrimination." *Perspectives on Politics* 18(4):1068–1087.
- Bauer, Nichole M. 2015a. "Emotional, Sensitive, and Unfit for Office? Gender Stereotype Activation and Support Female Candidates." *Political Psychology* 36(6):691–708.
- Bauer, Nichole M. 2015b. "Who Stereotypes Female Candidates? Identifying Individual Differences in Feminine Stereotype Reliance." *Politics, Groups, and Identities* 3(1):94–110.
- Bauer, Nichole M. 2017. "The Effects of Counterstereotypic Gender Strategies on Candidate Evaluations." *Political Psychology* 38(2):279–295.
- Bauer, Nichole M. 2018. "Untangling the Relationship Between Partisanship, Gender Stereotypes, and Support for Female Candidates." *Journal of Women, Politics and Policy* 39(1):1–25.
- Bauer, Nichole M. 2020a. "Shifting Standards: How Voters Evaluate the Qualifications of Female and Male Candidates." *Journal of Politics* 82(1):1–12.
- Bauer, Nichole M. 2020b. *The Qualifications Gap: Why Women Must Be Better than Men to Win Political Office*. Cambridge, UK: Cambridge University Press.
- Bauer, Nichole M., Laurel Harbridge Yong and Yanna Krupnikov. 2017. "Who is Punished? Conditions Affecting Voter Evaluations of Legislators Who Do Not Compromise." *Political Behavior* 39:279–300.
- Bauer, Nichole M., Nathan P. Kalmoe and Erica B. Russell. 2022. "Candidate Aggression and Gendered Voter Evaluations." *Political Psychology* 43(1):23–43.
- Bauer, Nichole M. and Tatum Taylor. 2022. "Selling them Short? Differences in News Coverage of Female and Male Candidate Qualifications." *Political Research Quarterly* (Forthcoming):1–15.

- Bear, Julia B. and Anita Williams Woolley. 2011. "The Role of Gender in Team Collaboration and Performance." *Interdisciplinary Science Reviews* 36(2):146–153.
- Beckwith, Karen. 2014. Plotting the Path from One to the Other: Women's Interests and Political Representation. In *Representation: The Case of Women's Interests*, ed. Michelle Taylor-Robinson and Maria Escobar-Lemmon. Oxford: Oxford University Press pp. 19–40.
- Benedetto, Giacomo and Simon Hix. 2007. "The Rejected, the Ejected, and the Dejected: Explaining Government Rebels in the 2001-2005 British House of Commons." *Comparative Political Studies* 40(7):755–780.
- Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society* 57(1):289–300.
- Benoit, Kenneth, Kevin Munger and Arthur Spirling. 2019. "Measuring and Explaining Political Sophistication Through Textual Complexity." *American Journal of Political Science* 63(2):491–508.
- Bergqvist, Christina, Elin Bjarnegård and Pär Zetterberg. 2018. "The Gendered Leeway: Male Privilege, Internal and External Mandates, and Gender-Equality Policy Change." *Politics, Groups, and Identities* 6(4):576–592.
- Bernhard, Rachel. 2022. "Wearing the Pants(suit)? Gendered Leadership Styles, Partisanship, and Candidate Evaluation in the 2016 U.S. Election." *Politics & Gender* 18(2):513–545.
- Bernhard, Rachel, Shauna Shames and Dawn Langan Teele. 2021. "To Emerge? Breadwinning, Motherhood, and Women's Decisions to Run for Office." *American Political Science Review* 115(2):379–394.
- Best, Deborah L. and John E. Williams. 1990. *Measuring Sex Stereotypes: A Thirty-Nation Study*. Beverly Hills, LA: Sage.
- Beutel, Ann M. and Margaret Mooney Marinin. 1995. "Gender and Values." *American Sociological Review* 60(3):436–448.
- Bevan, Shaun and Will Jennings. 2019. The UK Policy Agendas Project. In *Comparative Policy Agendas: Theory, Tools, Data*, ed. Frank R. Baumgartner, Christian Breunig and Emiliano Grossman. Oxford: Oxford University Press pp. 176–183.
- Bianchi, Suzanne, Laurent Lesnard, Tiziana Nazio and Sara Raley. 2014. "Gender and Time Allocation

- of Cohabiting and Married Women and Men in France, Italy, and the United States." *Demographic Research* 31(8):183–216.
- Biernat, Monica and Melvin Manis. 1994. "Shifting Standards and Stereotype-Based Judgments." *Journal of Personality and Social Psychology* 66(1):5–20.
- Biernat, Monica, Melvin Manis and Thomas E. Nelson. 1991. "Stereotypes and Standards of Judgment." *Journal of Personality and Social Psychology* 60(4):485–499.
- Bird, Karen. 2005. "Gendering Parliamentary Questions." *British Journal of Politics and International Relations* 7(1):353–370.
- Blair, Irene V. and Mahzarin R Banaji. 1996. "Automatic and Controlled Processes in Stereotype Priming." *Journal of Personality and Social Psychology* 70(6):1142–1163.
- Blankenship, Jane and Deborah C. Robson. 1995. "A 'Feminine Style' in Women's Political Discourse: An Exploratory Essay." *Communication Quarterly* 43(3):353–366.
- Blau, Francine D., Peter Brummund and Albert Yung-Hsu Liu. 2013. "Trends in Occupational Segregation by Gender 1970–2009: Adjusting for the Impact of Changes in the Occupational Coding System." *Demography* 50(2):471–492.
- Blei, David and John Lafferty. 2006. "Correlated Topic Models." *Advances in neural information processing systems* 18(147).
- Bligh, Michelle C. and Jeffrey C. Kohles. 2008. "Negotiating Gender Role Expectations: Rhetorical Leadership and Women in the US Senate." *Leadership* 4(4):381–402.
- Blumenau, Jack. 2021a. *House of Commons Parliamentary Debates, 1979-2019*. Colchester, Essex: UK Data Service.  
**URL:** [reshare.ukdataservice.ac.uk/854292/](https://reshare.ukdataservice.ac.uk/854292/)
- Blumenau, Jack. 2021b. "Online Activism and Dyadic Representation: Evidence from the UK E-Petition System." *Legislative Studies Quarterly* 46(4):889–920.
- Blumenau, Jack. 2021c. "The Effects of Female Leadership on Women's Voice in Political Debate." *British Journal of Political Science* 51(2):750–771.
- Blumenau, Jack and Benjamin E. Lauderdale. 2022a. "Liberals and Conservatives Rely on Very Similar Sets of Foundations When Comparing Moral Violations." *Working Paper* pp. 1–35.

- Blumenau, Jack and Benjamin E. Lauderdale. 2022b. "The Variable Persuasiveness of Political Rhetoric." *American Journal of Political Science* (Forthcoming):1–16.
- Blumenau, Jack, Fabio Wolkenstein and Christopher Wratil. 2022. "Measuring Multidimensional Representation." *Working Paper* .
- Blumenau, Jack and Roberta Damiani. 2021. The (Increasing) Discretion of MPs in Parliamentary Debate. In *The Politics of Legislative Debate*, ed. Hanna Bäck, Marc Debus and Jorge M. Fernandes. Oxford: Oxford University Press pp. 775–800.
- Bochel, Catherine and Jacqui Briggs. 2000. "Do Women Make A Difference?" *Politics* 20(2):63–68.
- Boussalis, Constantine, Gail Mcelroy and Miriam Sorace. 2019. "Exploring Female Legislators' Policy Agenda Using a Dynamic Topic Model." *Working Paper* .
- Boussalis, Constantine, Travis G. Coan, Mirya R. Holman and Stefan Müller. 2021. "Gender, Candidate Emotional Expression, and Voter Reactions During Televised Debates." *American Political Science Review* 115(4):1242–1257.
- Bratton, Kathleen A. and Leonard P. Ray. 2002. "Descriptive Representation, Policy Outcomes, and Municipal Day-Care Coverage in Norway." *American Journal of Political Science* 46(2):428.
- Brescoll, Victoria L. 2011. "Who Takes the Floor and Why: Gender, Power, and Volubility in Organizations." *Administrative Science Quarterly* 56(4):622–641.
- Brescoll, Victoria L. 2016. "Leading with their hearts? How gender stereotypes of emotion lead to biased evaluations of female leaders." *Leadership Quarterly* 27(3):415–428.
- Brescoll, Victoria L and Eric Luis Uhlmann. 2008. "Can an Angry Woman Get Ahead? Status Conferral, Gender, and Expression of Emotion in the Workplace." *Psychological Science* 19(3):268–275.
- Brooks, Deborah Jordan. 2011. "Testing the Double Standard for Candidate Emotionality: Voter Reactions to the Tears and Anger of Male and Female Politicians." *The Journal of Politics* 73(2):597–615.
- Brooks, Deborah Jordan. 2013. *He Runs, She Runs: Why Gender Stereotypes Do Not Harm Women Candidates*. New Jersey: Princeton University Press.
- Broughton, Sharon and Sonia Palmieri. 1999. "Gendered Contributions to Parliamentary Debates: The Case of Euthanasia." *Australian Journal of Political Science* 34(1):29–45.



- Broverman, Inge K., Susan R. Vogel, Frank E. Broverman, Frank E. Clarkson and Paul S. Rosenkrantz. 1972. "Sex-Role Stereotypes: A Current Appraisal." *Journal of Social Issues* 28:59-7.
- Burrell, Barbara C. 1994. *A Woman's Place Is in the House: Campaigning for Congress in the Feminist Era*. Michigan, U.S.: University of Michigan Press.
- Buser, Thomas, Muriel Niederle and Hessel Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *Quarterly Journal of Economics* 129(3):1409-1447.
- Bystrom, Dianne G., Terry A. Robertson and Mary Christine Banwart. 2001. "Framing the Fight: An Analysis of Media Coverage of Female and Male Candidates in Primary Races for Governor and U.S. Senate in 2000." *American Behavioral Scientist* 44(12):1999-2013.
- Cameron, Deborah and Sylvia Shaw. 2016. *Gender, Power and Political Speech*. Basingstoke: Palgrave Macmillan.
- Cameron, Deborah and Sylvia Shaw. 2020. "Constructing Women's "Different Voice": Gendered Mediation in the 2015 UK General Election." *Journal of Language and Politics* 19(1):144-159.
- Campbell, Rosie. 2012. "What Do We Really Know About Women Voters? Gender, Elections and Public Opinion." *The Political Quarterly* 83(4):703-710.
- Campbell, Rosie and Joni Lovenduski. 2015. "What Should MPs Do? Public and Parliamentarians' Views Compared." *Parliamentary Affairs* 68(4):690-708.
- Campbell, Rosie and Philip Cowley. 2014. "What Voters Want: Reactions to Candidate Characteristics in a Survey Experiment." *Political Studies* 62(4):745-765.
- Campbell, Rosie, Philip Cowley, Nick Vivyan and Markus Wagner. 2019a. "Legislator Dissent as a Aalence Signal." *British Journal of Political Science* 49(1):105-128.
- Campbell, Rosie, Philip Cowley, Nick Vivyan and Markus Wagner. 2019b. "Why Friends and Neighbors? Explaining the Electoral Appeal of Local Roots." *Journal of Politics* 81(3):937-951.
- Campbell, Rosie and Sarah Childs. 2010. "'Wags', 'Wives' and 'Mothers' ... But What About Women Politicians?" *Parliamentary Affairs* 63(4):760-777.
- Campbell, Rosie and Sarah Childs. 2014. "Parents in Parliament: 'Where's Mum?'" *The Political Quarterly* 85(4):487-492.

- Campbell, Rosie and Sarah Childs. 2015. "Conservatism, Feminisation and the Representation of Women in UK Politics." *British Politics* 10(1):148–168.
- Campbell, Rosie and Sarah Childs. 2017. The (Feminised) Contemporary Conservative Party. In *Rethinking right-wing women*, ed. Clarisse Berthezène and Julie V. Gottlieb. Manchester: Manchester University Press.
- Carli, Linda L. and Alice H. Eagly. 2017. Leadership and Gender. In *The Nature of Leadership*, ed. John Antonakis and David Day. Thousand Oaks, US: Sage pp. 244–271.
- Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76(1).
- Carroll, Susan J. and Kira Sanbonmatsu. 2013. *More Women Can Run: Gender and Pathways to the State Legislatures*. Oxford: Oxford University Press.
- Cassese, Erin C. and Mirya R. Holman. 2018. "Party and Gender Stereotypes in Campaign Attacks." *Political Behavior* 40:785–807.
- Catalano, Ana. 2009. "Women Acting for Women? An Analysis of Gender and Debate Participation in the British House of Commons 2005–2007." *Politics & Gender* 5(1):45–68.
- Cejka, Mary Ann and Alice H Eagly. 1999. "Gender-Stereotypic Images of Occupations Correspond to the Sex Segregation of Employment." *Personality and Social Psychology Bulletin* 25(4):413–423.
- Celis, Karen and Sarah Childs. 2012. "The Substantive Representation of Women: What to Do with Conservative Claims?" *Political Studies* 60(1):213–225.
- Celis, Karen, Sarah Childs, Johanna Kantola and Mona Lena Krook. 2014. "Constituting Women's Interests Through Representative Claims." *Politics and Gender* 10(2):149–174.
- Childs, Sarah. 2004a. "A Feminised Style of Politics? Women MPs in the House of Commons." *British Journal of Politics and International Relations* 6:3–19.
- Childs, Sarah. 2004b. *New Labour's Women MPs: Women Representing Women*. London: Routledge.
- Childs, Sarah. 2013. "Negotiating Gendered Institutions: Women's Parliamentary Friendships." *Politics & Gender* 9(1):127–151.
- Childs, Sarah. 2016. *The Good Parliament*. Bristol: University of Bristol.

- Childs, Sarah and Karen Celis. 2018. "Conservatism and Women's Political Representation." *Politics & Gender* 14(1):5–26.
- Childs, Sarah and Mona Lena Krook. 2009. "Analysing Women's Substantive Representation: From Critical Mass to Critical Actors." *Government and Opposition* 44(2):125–145.
- Childs, Sarah and Paul Webb. 2012. *Sex, Gender and the Conservative Party: From Iron Lady to Kitten Heals*. Basingstoke: Palgrave Macmillan.
- Cialdini, Robert B. and Noah J. Goldstein. 2004. "Social Influence: Compliance and Conformity." *Annual Review of Psychology* 55(1974):591–621.
- Claessen, Clint, Stefanie Bailer and Tomas Turner-Zwinkels. 2021. "The Winners of Legislative Mandate: An analysis of Post-Parliamentary Career Positions in Germany and the Netherlands." *European Journal of Political Research* 60(1):25–45.
- Clarke, Nick, Will Jennings, Jonathan Moss and Gerry Stoker. 2018. *The Good Politician: Folk Theories, Political Interaction, and the Rise of Anti-Politics*. Cambridge: Cambridge University Press.
- Clayton, Amanda, Cecilia Josefsson and Vibeke Wang. 2017. "Quotas and Women's Substantive Representation: Evidence from a Content Analysis of Ugandan Plenary Debates." *Politics and Gender* 13(2):276–304.
- Coates, Jennifer. 2015. *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language*. London: Routledge.
- Coffé, Hilde, Catherine Bolzendahl and Katia Schnellecke. 2019. "Parties, Issues, and Power: Women's Partisan Representation on German Parliamentary Committees." *European Journal of Politics and Gender* 2(2):257–281.
- Collet, Christian. 2008. "Minority Candidates, Alternative Media, and Multiethnic America: Deracialization or Toggling?" *Perspectives on Politics* 6(4):707–728.
- Collignon, Sofia, Rosie Campbell and Wolfgang Rüdig. 2022. "The Gendered Harassment of Parliamentary Candidates in the UK." *Political Quarterly* 93(1):32–38.
- Collignon, Sofia and Wolfgang Rüdig. 2020. "Harassment and Intimidation of Parliamentary Candidates in the United Kingdom." *Political Quarterly* 91(2):422–429.
- Collignon, Sofia and Wolfgang Rüdig. 2021. "Increasing the Cost of Female Representation? The Gen-

- dered Effects of Harassment, Abuse and Intimidation Towards Parliamentary Candidates in the UK." *Journal of Elections, Public Opinion and Parties* 31(4):429–449.
- Conroy, Meredith and Jon Green. 2020. "It Takes a Motive: Communal and Agentic Articulated Interest and Candidate Emergence." *Political Research Quarterly* 73(4):942–956.
- Cowley, Philip and Sarah Childs. 2003. "Too Spineless to Rebel? New Labour's Women MPs." *British Journal of Political Science* 33(2):345–365.
- Cuddy, Amy J.C., Susan T. Fiske and Peter Glick. 2008. "Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map." *Advances in Experimental Social Psychology* 40(7):61–149.
- Cutts, David, Matthew Goodwin, Oliver Heath and Paula Surridge. 2020. "Brexit, the 2019 General Election and the Realignment of British Politics." *Political Quarterly* 91(1):7–23.
- Dahlerup, Drude. 1988. "From a Small to a Large Minority: Women in Scandinavian Politics." *Scandinavian Political Studies* 11(4):275–298.
- Dahlerup, Drude. 2006. "The Story of the Theory of Critical Mass." *Politics & Gender* 2(4):511–522.
- Dahlerup, Drude. 2017. *Has Democracy Failed Women?* London, UK: Polity.
- Dalton, Russell J. and Martin P. Wattenberg. 2000. Unthinkable Democracy: Political Change in Advanced Industrial Democracies. In *Parties without Partisans*. Oxford: Oxford University Press pp. 3–19.
- Darcy, Robert, Susan Welch and Janet Clark. 1994. *Women, Elections & Representation*. Lincoln, Nebraska: University of Nebraska Press.
- Dasgupta, Nilanjana and Shaki Asgari. 2004. "Seeing is Believing: Exposure to Counterstereotypic Women Leaders and its Effect on the Malleability of Automatic Gender Stereotyping." *Journal of Experimental Social Psychology* 40(5):642–658.
- de Geus, Roosmarijn A., John R. McAndrews, Peter John Loewen and Aaron Martin. 2021. "Do Voters Judge the Performance of Female and Male Politicians Differently? Experimental Evidence from the United States and Australia." *Political Research Quarterly* 74(2):302–316.
- De Vries, Catherine E. and Sara B. Holbolt. 2020. *Political Entrepreneurs: The Rise of Challenger Parties in Europe*. New Jersey: Princeton University Press.

- Declercq, Eugene, Thomas L. Hurley and Norman R. Luttbeg. 1972. "Voting in American Presidential Elections, 1956-1972." *American Politics Quarterly* 3:222-243.
- Diekman, Amanda B and Alice H Eagly. 2000. "Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future." *Personality and Social Psychology Bulletin* 26(10):1171-1188.
- Diekman, Amanda B., Alice H. Eagly, Antonio Mladinic and Maria Cristina Ferreira. 2005. "Dynamic Stereotypes About Women and Men in Latin America and the United States." *Journal of Cross-Cultural Psychology* 36(2):209-226.
- Diekman, Amanda B. and Wind Goodfriend. 2006. "Rolling with the Changes: A Role Congruity Perspective on Gender Norms." *Psychology of Women Quarterly* 30(4):369-383.
- Dietrich, Bryce J., Matthew Hayes and Diana Z. O'Brien. 2019. "Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech." *American Political Science Review* 113(4):941-962.
- Ditonto, Tessa. 2017. "A High Bar or a Double Standard? Gender, Competence, and Information in Political Campaigns." *Political Behavior* 39:301-325.
- Ditonto, Tessa M., Allison J. Hamilton and David P. Redlawsk. 2014. "Gender Stereotypes, Information Search, and Voting Behavior in Political Campaigns." *Political Behavior* 36(2):335-358.
- Dittmar, Kelly, Kira Sanbonmatsu and Susan J. Carroll. 2018. *A Seat at the Table: Congresswomen's Perspectives on Why Their Presence Matters*. Oxford, UK: Oxford University Press.
- Dolan, Kathleen. 2004. *Voting for Women: How the Public Evaluates Women Candidates*. Boulder, U.S.: Westview Press.
- Dolan, Kathleen. 2014. "Gender Stereotypes, Candidate Evaluations, and Voting for Women Candidates: What Really Matters?" *Political Research Quarterly* 67(1):96-107.
- Donnelly, Kristin, Jean M. Twenge, Malissa A. Clark, Samia K. Shaikh, Angela Beiler-May and Nathan T. Carter. 2016. "Attitudes Toward Women's Work and Family Roles in the United States, 1976-2013." *Psychology of Women Quarterly* 40(1):41-54.
- Eagly, Alice H. 2013. *Sex Differences in Social Behavior: A Social-Role Interpretation*. New York: Psychology Press.
- Eagly, Alice H., Amanda B. Diekman, Mary C. Johannesen-Schmidt and Anne M. Koenig. 2004. "Gender

- Gaps in Sociopolitical Attitudes: A Social Psychological Analysis." *Journal of Personality and Social Psychology* 87(6):796–816.
- Eagly, Alice H, Christa Nater, David I Miller, Michèle Kaufmann and Sabine Sczesny. 2020. "Gender Stereotypes Have Changed: A Cross-Temporal Meta-Analysis of U.S. Public Opinion Polls From 1946 to 2018." *American Psychologist* 75(3):301–315.
- Eagly, Alice H. and Steven J. Karau. 2002. "Role Congruity Theory of Prejudice Toward Female Leaders." *Psychological Review* 109(3):573–598.
- Eagly, Alice H. and Valerie J. Steffen. 1984. "Gender Stereotypes Stem From the Distribution of Women and Men Into Social Roles." *Journal of Personality and Social Psychology* 46(4):735–754.
- Eagly, Alice H. and Wendy Wood. 2012. "Biosocial Construction of Sex Differences and Similarities in Behavior." *Advances in Experimental Social Psychology* 46:55–123.
- Eagly, Alice and Linda L. Carli. 2007. *Through the Labyrinth: The Truth about How Women Become Leaders*. Boston: Harvard Business Review.
- Eggers, Andrew C. and Arthur Spirling. 2017. "Incumbency Effects and the Strength of Party Preferences: Evidence from Multiparty Elections in the United Kingdom." *Journal of Politics* 79(3):903–920.
- Eggers, Andrew C., Nick Vivyan and Markus Wagner. 2017. "Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?" *Journal of Politics* 80(1):321–326.
- Elgie, Robert. 2015. *Studying Political Leadership: Foundations and Contending Accounts*. Basingstoke: Palgrave Macmillan.
- Enns-Jedenastik, Laurenz. 2017. "Campaigning on the Welfare State: The Impact of Gender and Gender Diversity." *Journal of European Social Policy* 27(3):215–228.
- Epitropaki, Olga and Robin Martin. 2004. "Implicit Leadership Theories in Applied Settings: Factor Structure, Generalizability, and Stability Over Time." *Journal of Applied Psychology* 89(2):293–310.
- Erikson, Josefina and Cecilia Josefsson. 2019. "The Legislature as a Gendered Workplace: Exploring Members of Parliament's Experiences of Working in the Swedish Parliament." *International Political Science Review* 40(2):197–214.
- Erikson, Josefina and Tània Verge. 2022. "Gender, Power and Privilege in the Parliamentary Workplace." *Parliamentary Affairs* 75(1):1–19.

- Escobar-Lemmon, Maria and Michelle M. Taylor-Robinson. 2005. "Women Ministers in Latin American Government: When, Where, and Why?" *American Journal of Political Science* 49(4):829–844.
- Everitt, Joanna, Lisa A. Best, Derek Gaudet and Joanna Everit. 2016. "Candidate Gender, Behavioral Style, and Willingness to Vote: Support for Female Candidates Depends on Conformity to Gender Norms." *American Behavioral Scientist* 60(14):1737–1755.
- Falk, Erika. 2008. *Women for President: Media Bias in Eight Campaigns*. Chicago: University of Chicago Press.
- Fenno, Richard F. 1978. *Home Style: House Members in Their Districts*. Boston, MA: Longman.
- Fernandes, Jorge M., Mariana Lopes da Fonseca and Miguel Won. 2021. "Closing the Gender Gap in Legislative Debates: The Role of Gender Quotas." *Political Behavior* (Forthcoming):1–25.
- Fieldhouse, Edward, Jane Green, Geoffrey Evans, Jonathan Mellon, Christopher Prosser, Hermann Schmitt and Cees van der Eijk. 2019. *Electoral Shocks: The Volatile Voter in a Turbulent World*. Oxford: Oxford University Press.
- Firth, John Rupert. 1957. *Studies in Linguistic Analysis*. New Jersey, US: Wiley-Blackwell.
- Fischbach, Andrea, Philipp W. Lichtenthaler and Nina Horstmann. 2015. "Leadership and Gender Stereotyping of Emotions: Think Manager - Think Male?" *Journal of Personnel Psychology* 14(3):153–162.
- Fiske, Susan T. and Laura E. Stevens. 1993. What's So Special About Sex? Gender Stereotyping and Discrimination. In *Gender Issues in Contemporary Society*, ed. Stuart Oskamp and Mark Costanzo. London: Sage Publications pp. 173–196.
- Fite, David, Marc Genest and Clyde Wilcox. 1990. "Foreign Policy Attitudes." *American Politics Quarterly* 18(4):492–513.
- Foley, Michael. 2013. *Political Leadership: Themes, Contexts and Critiques*. Oxford: Oxford University Press.
- Fox, Richard L. and Jennifer L. Lawless. 2004. "Entering the Arena? Gender and the Decision to Run for Office." *American Journal of Political Science* 48(2):264–280.
- Fox, Richard L. and Zoe M. Oxley. 2003. "Gender Stereotyping State Executive Elections: Candidate Selection and Success." *Journal of Politics* 65(3):833–50.
- Franceschet, Susan. 2011. Gendered Institutions and Women's Substantive Representation: Female

- Legislators in Argentina and Chile. In *Gender, Politics and Institutions: Towards a Feminist Institutionalism*, ed. Mona Lena Krook and Fiona Mackay. Basingstoke: Palgrave Macmillan pp. 58–78.
- Francis, Becky, Barbara Read, Lindsay Melling and Jocelyn Robson. 2003. "University Lecturers' Perceptions of Gender and Undergraduate Writing." *British Journal of Sociology of Education* 24(3):357–373.
- Fulton, Sarah A. 2012. "Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65(2):303–314.
- Garzia, Diego. 2011. "The Personalization of Politics in Western Democracies: Causes and Consequences on Leader-Follower Relationships." *Leadership Quarterly* 22(4):697–709.
- Gelman, Andrew and Iain Pardoe. 2006. "Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models." *Technometrics* 48(2):241–251.
- Gelman, Andrew, Jennifer Hill and Masanao Yajima. 2012. "Why We (Usually) Don't Have to Worry About Multiple Comparisons." *Journal of Research on Educational Effectiveness* 5(2):189–211.
- Gidengil, Elisabeth and Joanna Everitt. 2003. "Talking Tough: Gender and Reported Speech in Campaign News Coverage." *Political Communication* 20(3):209–232.
- Gill, Michael J. 2004. "When Information Does Not Deter Stereotyping: Prescriptive Stereotyping Can Foster Bias Under Conditions That Deter Descriptive Stereotyping." *Journal of Experimental Social Psychology* 40(5):619–632.
- Gleason, Shane A. 2020. "Beyond Mere Presence: Gender Norms in Oral Arguments at the U.S. Supreme Court." *Political Research Quarterly* 73(3):596–608.
- Gleason, Shane A., Jennifer J. Jones and Jessica Rae McBean. 2019. "The Role of Gender Norms in Judicial Decision-Making at the U.S. Supreme Court: The Case of Male and Female Justices." *American Politics Research* 47(3):494–529.
- Gomard, Kirsten and Anne Krogstad. 2001. *Instead of the Ideal Debate: Doing Politics and Doing Gender in Nordic Political Campaign Discourse*. Aarhus, Denmark: Aarhus University Press.
- Goodwin, Mark, Stephen Holden Bates and Stephen McKay. 2021. "Electing to Do Women's Work? Gendered Divisions of Labor in U.K. Select Committees, 1979–2016." *Politics and Gender* 17(4):607–639.
- Green, Jane and Will Jennings. 2017. *The Politics of Competence: Parties, Public Opinion and Voters*. Cambridge: Cambridge University Press.



- Greene, Zachary and Diana Z. O'Brien. 2016. "Diverse Parties, Diverse Agendas? Female Politicians and the Parliamentary Party's Role in Platform Formation." *European Journal of Political Research* 55(3):435-453.
- Grey, Sandra. 2002. "Does Size Matter? Critical Mass and New Zealand's Women MPs." *Parliamentary Affairs* 55:19-29.
- Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21:267-297.
- Hansard Society. 2014. *Tuned in or Turned off? Public attitudes to Prime Minister's Questions*. London: Hansard Society.
- Hargrave, Lotte. 2021. "A Double Standard? Gender Bias in Voters' Perceptions of Political Arguments." *EGAP* (ID: 20210908AB).  
**URL:** [osf.io/bfds7](https://osf.io/bfds7)
- Hargrave, Lotte and Jessica C. Smith. 2022. "Working Hard or Hardly Working? Gender and Voter Evaluations of Legislator Productivity." *Working Paper* pp. 1-33.
- Hargrave, Lotte and Tone Langengen. 2021. "The Gendered Debate: Do Men and Women Communicate Differently in the House of Commons?" *Politics & Gender* 17(4):580-606.
- Harmer, Emily. 2021. *Women, Media, and Elections: Representation and Marginalization in British Politics*. Bristol: Bristol University Press.
- Haselmayer, Martin, Sarah C. Dingler and Marcelo Jenny. 2021. "How Women Shape Negativity in Parliamentary Speeches—A Sentiment Analysis of Debates in the Austrian Parliament." *Parliamentary Affairs* (Forthcoming):1-20.
- Hawkesworth, Mary. 2003. "Congressional Enactments of Race-Gender: Toward a Theory of Raced-Gendered Institutions." *American Political Science Review* 97(4):529-550.
- Hawkesworth, Mary. 2005. "Engendering Political Science: An Immodest Proposal." *Politics & Gender* 1(1):141-156.
- Hayes, Danny. 2011. "When Gender and Party Collide: Stereotyping in Candidate Trait Attribution." *Politics and Gender* 7(2):133-165.

- Heilman, M. E. 2001. "Description and Prescription: How Gender Stereotypes Prevent Women's Ascent up the Organizational Ladder." *Journal of Social Issues* 57(4):657-674.
- Heldman, Caroline, Susan J. Carroll and Stephanie Olson. 2005. "'She Brought Only A Skirt': Print Media Coverage of Elizabeth Dole's Bid for the Republican Presidential Nomination." *Political Communication* 22(3):315-335.
- Herrnson, Paul S., J. Celeste Lay and Atiya Kai Stokes. 2003. "Women Running "as Women": Candidate Gender, Campaign Issues, and Voter-Targeting Strategies." *The Journal of Politics* 65(1):244-255.
- Höhmann, Daniel. 2020. "When Do Men Represent Women's Interests in Parliament? How the Presence of Women in Parliament Affects the Legislative Behavior of Male Politicians." *Swiss Political Science Review* 26(1):31-50.
- Höhmann, Daniel and Mary Nugent. 2021. "Male MPs, electoral vulnerability and the substantive representation of women's interests." *European Journal of Political Research* (Forthcoming).
- Holman, Mirya R. and Anna Mahoney. 2018. "Stop, Collaborate, and Listen: Women's Collaboration in US State Legislatures." *Legislative Studies Quarterly* 43(2):179-206.
- Holman, Mirya R., Anna Mahoney and Emma Hurler. 2021. "Let's Work Together: Bill Success via Women's Cosponsorship in U.S. State Legislatures." *Political Research Quarterly* .
- Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2011. "Sex, Stereotypes, and Security: A Study of the Effects of Terrorist Threat on Assessments of Female Leadership." *Journal of Women, Politics & Policy* 32(3):173-192.
- Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2016. "Terrorist Threat, Male Stereotypes, and Candidate Evaluations." *Political Research Quarterly* 69(1):134-147.
- Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2017. "Can Experience Overcome Stereotypes in Times of Terror Threat?" *Research and Politics* 4(1):1-7.
- Holmes, Janet and Miriam Meyerhoff. 2003. *The Handbook of Language and Gender*. Oxford: Blackwell Publishing.
- Hornikx, Jos. 2018. "Combining Anecdotal and Statistical Evidence in Real-Life Discourse: Comprehension and Persuasiveness." *Discourse Processes* 55(3):324-336.
- Huddy, Leonie, Erin C. Cassese and Mary-Kate Lizotte. 2008. *Gender, Public Opinion, and Political Rea-*

- soning. In *Political Women and American Democracy*, ed. Christina Wolbrecht, Karen Beckwith and Lisa Baldez. Cambridge: Cambridge University Press pp. 31–49.
- Huddy, Leonie and Nayda Terkildsen. 1993a. “Gender Stereotypes and the Perception of Male and Female Candidates.” *American Journal of Political Science* 37(1):119–147.
- Huddy, Leonie and Nayda Terkildsen. 1993b. “The Consequences of Gender Stereotypes for Women Candidates at Different Levels and Types of Office.” *Political Research Quarterly* 46(3):503.
- Humphreys, Macartan, William A. Masters and Martin E. Sandbu. 2006. “The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe.” *World Politics* 58(4):583–622.
- Hurwitz, Jon and Shannon Smitley. 1998. “Gender Differences and Crime and Punishment.” *Political Research Quarterly* 51(1):89–115.
- Inglehart, Ronald and Pippa Norris. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge, UK: Cambridge University Press.
- IPU. 2022. “Women in National Parliaments.”  
**URL:** <http://archive.ipu.org/wmn-e/world.htm>
- Iversen, Torben and Frances Rosenbluth. 2010. *Women, Work, and Politics: The Political Economy of Gender Inequality*. New Haven: Yale University Press.
- Jacobi, Tonja and Dylan Schweers. 2017. “Justice, Interrupted: The Effect of Gender, Ideology, and Seniority at Supreme Court Oral Arguments.” *Virginia Law Review* .
- Jalalzai, Farida. 2016. “Shattered Not Cracked: The Effect of Women’s Executive Leadership.” *Journal of Women, Politics & Policy* 37(4):439–463.
- Jamieson, Kathleen. 1995. *Beyond the Double Bind: Women and Leadership*. Oxford: Oxford University Press.
- Johnson, Danette Ifert. 2005. “Feminine Style in Presidential Debate Discourse, 1960–2000.” *Communication Quarterly* 53(1):3–20.
- Jones, Jennifer J. 2016. “Talk “Like a Man”: The Linguistic Styles of Hillary Clinton, 1992–2013.” *Perspectives on Politics* 14(3):625–642.

- Kahn, Kim Fridkin. 1992. "Does Being Male Help? An Investigation of the Effects of Candidate Gender and Campaign Coverage on Evaluations of U.S. Senate Candidates." *The Journal of Politics* 54(2):497–517.
- Kahn, Kim Fridkin. 1994. "The Distorted Mirror: Press Coverage of Women Candidates for Statewide Office." *The Journal of Politics* 56(1):154–173.
- Kahn, Kim Fridkin. 1996. *The Political Consequences of Being a Woman: How Stereotypes Influence the Conduct and Consequences of Political Campaigns*. New York, US: Columbia University Press.
- Kam, Christopher J. 2009. *Party Discipline and Parliamentary Politics*. Cambridge: Cambridge University Press.
- Kanthak, Kristin and Jonathan Woon. 2015. "Women Don't Run? Election Aversion and Candidate Entry." *American Journal of Political Science* 59(3):595–612.
- Karpowitz, Christopher F. and Tali Mendelberg. 2014. *The Silent Sex: Gender, Deliberation & Institutions*. Princeton: Princeton University Press.
- Karpowitz, Christopher F., Tali Mendelberg and Lauren Mattioli. 2015a. "How Group Forces Demonstrate the Malleability of Gendered Behavior." *Politics, Groups, and Identities* 3(1):203–208.
- Karpowitz, Christopher F., Tali Mendelberg and Lauren Mattioli. 2015b. "Why Women's Numbers Elevate Women's Influence, and When They Do Not: Rules, Norms, and Authority in Political Discussion." *Politics, Groups, and Identities* 3(1):149–177.
- Karvonen, Lauri, Göran Djupsund and Tom Carlson. 1995. Political Language. In *Women in Nordic Politics: Closing the Gap*, ed. L Karvonen and P Selle. Farnham, UK: Ashgate Publishing pp. 343–382.
- Kathlene, Lyn. 1994. "Power and Influence in State Legislative Policymaking: The Interaction of Gender and Position in Committee Hearing Debates." *American Political Science Review* 88(3):560–576.
- Kathlene, Lyn. 2001. Words That Matter: Women's Voice and Institutional Bias in Public Policy Formation. In *The Impact of Women in Public Office*, ed. Susan J. Carroll. Indiana, US: Indiana University Press pp. 22–48.
- Kincaid, J. Peter, Robert P. Fishburne Jr, Richard L. Rogers and Brad S. Chissom. 1975. Derivation Of New Readability Formulas (Automated Readability Index, Fog Count And Flesch Reading Ease Formula) For Navy Enlisted Personnel. Technical report Institute for Simulation and Training.

- Kittilson, Miki C. and Leslie A. Schwindt-Bayer. 2012. *The Gendered Effects of Electoral Institutions: Political Engagement and Participation*. Oxford: Oxford University Press.
- Kittilson, Miki Caul. 2011. "Women, Parties and Platforms in Post-Industrial Democracies." *Party Politics* 17(1):66–92.
- Klofstad, Casey A., Rindy C. Anderson and Stephen Nowicki. 2015. "Perceptions of Competence, Strength, and Age Influence Voters to Select Leaders with Lower-Pitched Voices." *PLoS ONE* 10(8):1–14.
- Knobloch-Westerwick, Silvia, Carroll J. Glynn and Michael Hoge. 2013. "The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest." *Science Communication* 35(5):603–625.
- Koch, Jeffrey W. 2000. "Do Citizens Apply Gender Stereotypes to Infer Candidates' Ideological Orientations?" *The Journal of Politics* 62(2):414–429.
- Koch, Jeffrey W. 2002. "Gender Stereotypes and Citizens' Impressions of House Candidates' Ideological Orientations." *American Journal of Political Science* 46(2):453–462.
- Koenig, Anne M., Alice H. Eagly, Abigail A. Mitchell and Tiina Ristikari. 2011. "Are Leader Stereotypes Masculine? A Meta-Analysis of Three Research Paradigms." *Psychological Bulletin* 137(4):616–642.
- Krook, Mona Lena and Diana Z. O'Brien. 2012. "All the President's Men? The Appointment of Female Cabinet Ministers Worldwide." *The Journal of Politics* 74(3):840–855.
- Krook, Mona Lena and Fiona Mackay. 2011. *Gender, Politics and Institutions: Towards a Feminist Institutionalism*. Basingstoke, UK: Palgrave Macmillan.
- Krupnikov, Yanna and Nichole M. Bauer. 2014. "The Relationship Between Campaign Negativity, Gender and Campaign Context." *Political Behavior* 36:167–188.
- Lamprinakou, Chrysa, Marco Morucci, Rosie Campbell and Jennifer Van Heerde-Hudson. 2017. "All Change in the House? The Profile of Candidates and MPs in the 2015 British General Election." *Parliamentary Affairs* 70:207–232.
- Lawless, Jennifer L. 2004. "Women, War, and Earning Elections: Gender Stereotyping in the Post-September 11th Era." *Political Research Quarterly* 57(3):479–490.
- Lawless, Jennifer L. and Richard L. Fox. 2005. *It Takes A Candidate: Why Women Don't Run for Office*. Cambridge, UK: Cambridge University Press.

- Lawless, Jennifer L. and Richard L. Fox. 2010. *It Still Takes a Candidate: Why Women Don't Run for Office*. Cambridge: Cambridge University Press.
- Lawless, Jennifer L. and Richard L. Fox. 2012. *Men Rule: The Continued Under-Representation of Women in U.S. Politics*. Washington, DC: Women & Politics Institute.
- Lawless, Jennifer L., Sean M. Theriault and Samantha Guthrie. 2018. "Nice Girls? Sex, Collegiality, and Bipartisan Cooperation in the US Congress." *The Journal of Politics* 80(4):1268–1282.
- Lazarus, Jeffrey and Amy Steigerwalt. 2018. *Gendered Vulnerability: How Women Work Harder to Stay in Office*. Michigan, U.S.: University of Michigan Press.
- Lazarus, Jeffrey, Amy Steigerwalt and Micayla Clark. 2022. "Time Spent in the House: Gender and the Political Careers of U.S. House Members." *Politics & Gender* (Forthcoming):1–36.
- Leeper, Campbell and Melanie M Ayres. 2007. "A Meta-Analytic Review of Gender Variations in Adults' Language Use: Talkativeness, Affiliative Speech, and Assertive Speech." *Personality and Social Psychology Review* 11(4):328–363.
- Lijphart, Arend. 2012. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.
- Lin, Nick and Moritz Osnabrügge. 2018. "Making Comprehensible Speeches When Your Constituents Need It." *Research and Politics* 5(3):30–32.
- Lippa, Richard A., Kathleen Preston and John Penner. 2014. "Women's Representation in 60 Occupations from 1972 to 2010: More Women in High-Status Jobs, Few Women in Things-Oriented Jobs." *PLoS ONE* 9(5):1–8.
- Lovenduski, Joni. 2012. "Feminising British Politics." *The Political Quarterly* 83(4):697–702.
- Mackay, Fiona. 2008. "'Thick' Conceptions of Substantive Representation: Women, Gender and Political Institutions." *Representation* 44(2):125–139.
- Mackay, Fiona, Meryl Kenny and Louise Chappell. 2010. "New Institutionalism Through a Gender Lens: Towards a Feminist Institutionalism?" *International Political Science Review* 31(5):573–588.
- Mansbridge, Jane. 2003. "Rethinking Representation." *American Political Science Review* 97(4):515–528.

- Martindale, Colin. 1990. *The Clockwork Muse: The Predictability of Artistic Change*. New York: Basic Books.
- Matland, Richard E. and Donley T. Studlar. 1996. "The Contagion of Women Candidates in Single-Member District and Proportional Representation Electoral Systems: Canada and Norway." *The Journal of Politics* 58(3):707-733.
- Mattei, Laura R. Winsky. 1998. "Gender and Power in American Legislative Discourse." *The Journal of Politics* 60(2):440-461.
- McDermott, Monika L. 1997. "Voting Cues in Low-Information Elections: Candidate Gender as a Social Information Variable in Contemporary United States Elections." *American Journal of Political Science* 41(1):270-283.
- Meguid, Bonnie M., Hilde Coffé, Ana Catalano Weeks and Miki Caul Kittilson. 2022. "The New Defenders of Gender Equality: When Do Populist Radical Right Parties Incorporate Women's Interests?" *Working Paper* .
- Mendelberg, Tali, Christopher F. Karpowitz and J. Baxter Oliphant. 2014. "Gender Inequality in Deliberation: Unpacking the Black Box of Interaction." *Perspectives on Politics* 12(1):18-44.
- Mendelberg, Tali, Christopher F. Karpowitz and Nicholas Goedert. 2013. "Does Descriptive Representation Facilitate Women's Distinctive Voice? How Gender Composition and Decision Rules Affect Deliberation." *American Journal of Political Science* 58(2):291-306.
- Meyrowitz, Joshua. 1985. *No Sense of Place: The Impact of Electronic Media on Social Behavior*. Oxford: Oxford University Press.
- Mikolov, Tomas, Kai Chen, Greg Corrado and Jeffrey Dean. 2013. *Efficient Estimation of Word Representations in Vector Space*. Working Paper.
- Miquel, Gerard Padró I and James M. Snyder. 2006. "Legislative Effectiveness and Legislative Careers." *Legislative Studies Quarterly* 31(3):347-381.
- Mo, Cecilia Hyunjung. 2015. "The Consequences of Explicit and Implicit Gender Attitudes and Candidate Quality in the Calculations of Voters." *Political Behavior* 37:357-395.
- Mutz, Diana C. 2011. *Population-Based Survey Experiments*. Princeton: Princeton University Press.

- Norris, Pippa. 1996. "Women Politicians: Transforming Westminster?" *Parliamentary Affairs* 49(1):89–102.
- Norris, Pippa and Joni Lovenduski. 1995. *Political Recruitment: Gender, Race and Class in the British Parliament*. Cambridge, UK: Cambridge University Press.
- Northouse, Peter G. 2016. *Leadership: Theory and Practice*. London: SAGE Publications.
- O'Brien, Diana Z. and Johanna Rickne. 2016. "Gender Quotas and Women's Political Leadership." *American Political Science Review* 110(1):112–126.
- Offermann, Lynn R., John K. Kennedy Jr. and Philip W. Wirtz. 1994. "Implicit Leadership Theories: Content, Structure, and Generalizability." *Leadership Quarterly* 5(1):43–58.
- O'Grady, Tom. 2019. "Careerists Versus Coal-Miners: Welfare Reforms and the Substantive Representation of Social Groups in the British Labour Party." *Comparative Political Studies* 52(4):544–578.
- O'Grady, Tom. 2022. *The Transformation of British Welfare Policy: Politics, Discourse, and Public Opinion*. Oxford: Oxford University Press.
- Okimoto, Tyler G. and Victoria L. Brescoll. 2010. "The Price of Power: Power Seeking and Backlash Against Female Politicians." *Personality and Social Psychology Bulletin* 36(7):923–936.
- Oritiz-Ospina, Esteban and Sandra Tzvetkova. 2017. Working Women: Key Facts and Trends in Female Labor Force Participation. Technical report Our World in Data.
- Osnabrügge, Moritz, Sara B. Hobolt and Toni Rodon. 2021. "Playing to the Gallery: Emotive Rhetoric in Parliaments." *American Political Science Review* 115(3):885–899.
- Page, Benjamin I. and Robert Y. Shapiro. 2010. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago, US: University of Chicago Press.
- Pasek, Josh and Jon A. Krosnick. 2010. Optimizing Survey Questionnaire Design in Political Science: Insights from Psychology. In *The Oxford Handbook of American Elections and Political Behavior*, ed. Jan E. Leighley. Oxford, UK: Oxford University Press pp. 27–50.
- Pearson, Kathryn and Logan Dancey. 2011a. "Elevating Women's Voices in Congress: Speech Participation in the House of Representatives." *Political Research Quarterly* 64(4):910–923.
- Pearson, Kathryn and Logan Dancey. 2011b. "Speaking for the Underrepresented in the House of Representatives: Voicing Women's Interests in a Partisan Era." *Politics and Gender* 7(4):493–519.



- Pennebaker, James W., Ryan L. Boyd, Kayla Jordan and Kate Blackburn. 2015. *The Development and Psychometric Properties of LIWC2015*. Austin, US: University of Texas at Austin.
- Pennington, Jeffrey, Richard Socher and Christopher D Manning. 2014. GloVe: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. pp. 1532–1543.
- Petrocik, John R., William L. Benoit and Glenn J. Hansen. 2003. "Issue Ownership and Presidential Campaigning , 1952-2000." *Political Science Quarterly* 118(4):599–626.
- Phillips, Anne. 1995. *The Politics of Presence*. Oxford, UK: Oxford University Press.
- Piscopo, Jennifer M. 2011. "Rethinking Descriptive Representation: Rendering Women in Legislative Debates." *Parliamentary Affairs* 64(3):448–472.
- Piscopo, Jennifer M. 2020. "Women Leaders and Pandemic Performance: A Spurious Correlation." *Politics & Gender* .
- Pitkin, Hanna. 1967. *The Concept of Representation*. Los Angeles: University of California Press.
- Profeta, Paola and Eleanor Florence Woodhouse. 2022. "Electoral Rules, Women's Representation and the Qualification of Politicians." *Comparative Political Studies* (Forthcoming):1–30.
- Proksch, Sven-Oliver and Jonathan B. Slapin. 2015. *The Politics of Parliamentary Debate: Parties, Rebels, and Representation*. Cambridge: Cambridge University Press.
- Proksch, Sven Oliver, Will Lowe, Jens Wäckerle and Stuart Soroka. 2019. "Multilingual Sentiment Analysis: A New Approach to Measuring Conflict in Legislative Speeches." *Legislative Studies Quarterly* 44(1):97–131.
- Puwar, Nirmal. 2004. *Space Invaders: Race, Gender and Bodies Out of Place*. Oxford, UK: Berg.
- Queneau, Hervé. 2006. "Is the Long-Term Reduction in Occupational Sex Segregation Still Continuing in the United States?" *The Social Science Journal* 43:681–688.
- Reher, Stefanie. 2022. "Do Disabled Candidates Represent Disabled Citizens?" *British Journal of Political Science* 52(2):520–534.
- Reingold, Beth. 2000. *Representing Women: Sex, Gender, and Legislative Behavior in Arizona and California*. North Carolina, US: The University of North Carolina Press.

- Renwick, Alan, Ben Lauderdale, Meg Russell and James Cleaver. 2022. *What Kind of Democracy Do People Want?* London: The Constitution Unit.
- Rice, Douglas R. and Christopher Zorn. 2021. "Corpus-Based Dictionaries for Sentiment Analysis of Specialized Vocabularies." *Political Science Research and Methods* 9(1):20–35.
- Roberts, Margaret E., Brandon M. Stewart and Dustin Tingley. 2014. "stm: R Package for Structural Topic." *Journal of Statistical Software* 10(2):1–40.
- Robertson, Terry, Kristin Froemling, Scott Wells and Shannon Mccraw. 1999. "Sex, Lies, and Videotape: An Analysis of Gender in Campaign Advertisements." *Communication Quarterly* 47(3):333–341.
- Rodriguez, Pedro L. and Arthur Spirling. 2022. "Word Embeddings: What Works, What Doesn't, and How to Tell the Difference for Applied Research." *Journal of Politics* 84(1):101–115.
- Ross, Karen. 2002. "Women's Place in 'Male' Space: Gender and Effect in Parliamentary Contexts." *Parliamentary Affairs* 55:189–201.
- Ross, Karen, Elizabeth Evans, Lisa Harrison, Mary Shears and Khursheed Wadia. 2013. "The Gender of News and News of Gender: A Study of Sex, Politics, and Press Coverage of the 2010 British General Election." *International Journal of Press/Politics* 18(1):3–20.
- Rucker, Derek D., Adam D. Galinsky and Joe C. Magee. 2018. "The Agentic–Communal Model of Advantage and Disadvantage: How Inequality Produces Similarities in the Psychology of Power, Social Class, Gender, and Race." *Advances in Experimental Social Psychology* 58:71–125.
- Rudman, Laurie A. and Stephanie A. Goodwin. 2004. "Gender Differences in Automatic In-Group Bias: Why Do Women Like Women More Than Men Like Men?" *Journal of Personality and Social Psychology* 87(4):494–509.
- Rush, Michael and Philip Giddings. 2011. *Parliamentary Socialisation*. London: Palgrave Macmillan.
- Russell, Meg. 2005. *Must Politics Disappoint?* London: Fabian Society.
- Saha, Sparsha and Ana Catalano Weeks. 2022. "Ambitious Women: Gender and Voter Perceptions of Candidate Ambition." *Political Behavior* 44:779–805.
- Salmela, Mikko and Christian Von Scheve. 2018. "Emotional Dynamics of Right-and Left-wing Political Populism." *Humanity & Society* 42(4):434–454.

- Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science* 46(1):20–34.
- Sanbonmatsu, Kira and Kathleen Dolan. 2009. "Do Gender Stereotypes Transcend Party?" *Political Research Quarterly* 62(3):485–494.
- Sayer, Liana C., Suzanne M. Bianchi and John P. Robinson. 2004. "Are Parents Investing Less in Children? Trends in Mothers' and Fathers' Time with Children." *American Journal of Sociology* 110(1):1–43.
- Schneider, Monica C. 2014a. "Gender-Based Strategies on Candidate Websites." *Journal of Political Marketing* 13(4):264–290.
- Schneider, Monica C. 2014b. "The Effects of Gender-Bending on Candidate Evaluations." *Journal of Women, Politics and Policy* 35(1):55–77.
- Schneider, Monica C. and Angela L. Bos. 2014. "Measuring Stereotypes of Female Politicians." *Political Psychology* 35(2):245–266.
- Schneider, Monica C. and Angela L. Bos. 2019. "The Application of Social Role Theory to the Study of Gender in Politics." *Political Psychology* 40(1):173–213.
- Schwarz, Susanne and Alexander Coppock. 2022. "What Have We Learned about Gender from Candidate Choice Experiments? A Meta-Analysis of Sixty-Seven Factorial Survey Experiments." *Journal of Politics* 84(2).
- Schwindt-Bayer, Leslie A. 2006. "Still Supermadres? Gender and the Policy Priorities of Latin American Legislators." *American Journal of Political Science* 50(3):570–585.
- Sczesny, Sabine, Christa Nater and Alice H. Eagly. 2019. Agency and Communion: Their Implications for Gender Stereotypes and Gender Identities. In *Agency and Communion in Social Psychology*, ed. Andrea E. Abele and Bogdan Wojciszke. New York, US: Routledge pp. 103–116.
- Searles, Kathleen, Erika Franklin Fowler, Travis N. Ridout, Patricia Strach and Katherine Zuber. 2020. "The Effects of Men's and Women's Voices in Political Advertising." *Journal of Political Marketing* 19(3):301–329.
- Seguino, Stephanie. 2007. "PlusÇa Change? Evidence on global trends in gender norms and stereotypes." *Feminist Economics* 13(2):1–28.

- Sendén, Marie Gustafsson, Amanda Klysing, Anna Lindqvist and Emma Aurora Renström. 2019. "The (Not So) Changing Man: Dynamic Gender Stereotypes in Sweden." *Frontiers in Psychology* 10(JAN):1–17.
- Sevi, Semra. 2022. "Is Incumbency Advantage Gendered?" *Legislative Studies Quarterly* (Forthcoming):1–19.
- Shames, Shauna L. 2017. *Out of the Running: Why Millennials Reject Political Careers and Why it Matters*. New York: New York University Press.
- Shapiro, Robert Y. 1986. "Gender Differences in Policy Preferences: A Summary of Trends from the 1960s to the 1980s." *Public Opinion Quarterly* 50:42–61.
- Shaw, Sylvia. 2000. "Language, Gender and Floor Apportionment in Political Debates." *Discourse & Society* 11(3):401–418.
- Shaw, Sylvia. 2002. *Language and Gender in Political Debates in the House of Commons* PhD thesis University of London.
- Shaw, Sylvia. 2006. *Governed by the Rules?: The Female Voice in Parliamentary Debates*. In *Speaking Out: The Female Voice in Public Contexts*, ed. Judith Baxter. Basingstoke, UK: Palgrave Macmillan pp. 81–95.
- Shaw, Sylvia. 2020. *Women, Language and Politics*. Cambridge: Cambridge University Press.
- Shorrocks, Rosalind. 2018. "A Feminist Generation? Cohort Change in Gender-Role Attitudes and the Second-Wave Feminist Movement." *International Journal of Public Opinion Research* 30(1):125–145.
- Skjeie, Hege. 1991. "The Rhetoric of Difference: On Women's Inclusion into Political Elites." *Politics & Society* 19(2):233–263.
- Smith, Jessica C. 2019. *Gendering Political Leadership: A Case Study of the UK , Examining Media and Voter Perceptions* PhD thesis Birkbeck, University of London.
- Sobolewska, Maria, Rebecca McKee and Rosie Campbell. 2018. "Explaining Motivation to Represent: How Does Descriptive Representation Lead to Substantive Representation of Racial and Ethnic Minorities?" *West European Politics* 41(6):1237–1261.
- Sones, Boni, Margaret Moran and Joni Lovenduski. 2005. *Women in Parliament: The New Suffragettes*. London, UK: Politico's Publishing.
- Spence, Janet T., Robert Helmreich and Joy Stapp. 1974. "The Personal Attributes Questionnaire: A

- Measure of Sex Role Stereotypes and Masculinity-Femininity." *Catalog of Selected Documents in Psychology* 4:43-44.
- Strand, Elizabeth A. 1999. "Uncovering the Role of Gender Stereotypes in Speech Perception." *Journal of Language and Social Psychology* 18(1):86-99.
- Strøm, Kaare. 1997. "Rules, Reasons and Routines: Legislative Roles in Parliamentary Democracies." *The Journal of Legislative Studies* 3(1):155-174.
- Swers, Michele L. 2002. *The Difference Women Make: The Policy Impact of Women in Congress*. Chicago: University of Chicago Press.
- Swift, Clint S. and Kathryn Van der Molen. 2021. "Marginalization and Mobilization: The Roots of Female Legislators' Collaborative Advantage in the States." *State Politics & Policy Quarterly* 21(4):355-379.
- Taylor, Eleanor Attar and Jacqueline Scott. 2018. "Gender: New consensus or continuing battleground?" *British Social Attitudes* 35 pp. 56-85.
- Teele, Dawn Langan. 2018. *Forging the Franchise: The Political Origins of Women's Vote*. New Jersey, US: Princeton University Press.
- Teele, Dawn Langan, Joshua Kalla and Frances Rosenbluth. 2018. "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* 112(3):525-541.
- Thelander, Kerstin. 1986. *Politikersprak i konsperspektiv*. Malmo: Liber Forlag.
- Thomas, Sue. 1994. *How Women Legislate*. Oxford, UK: Oxford University Press.
- Togebly, Lise. 1994. "The Gender Gap in Foreign Policy Attitudes." *Journal of Peace Research* 31(4):375-392.
- Tolleson-Rinehart, Sue. 2001. Do Women Leaders Make a Difference? Substance, Style, and Perceptions. In *The Impact of Women in Public Office*, ed. Susan J Carroll. Indiana: Indiana University Press pp. 149-165.
- Twenge, Jean M. 1997a. "Attitudes Toward Women, 1970-1995: A Meta-Analysis." *Psychology of Women Quarterly* 21(1):35-51.
- Twenge, Jean M. 1997b. "Changes in Masculine and Feminine Traits Over Time: A Meta-Analysis." *Sex Roles* 36:305-325.

- Twenge, Jean M. 2001. "Changes in Women's Assertiveness in Response to Status and Roles: A Cross-Temporal Meta-Analysis, 1931-1993." *Journal of Personality and Social Psychology* 81(1):133-145.
- Utych, Stephen M. 2021. "Speaking Style and Candidate Evaluations." *Politics, Groups, and Identities* 9(3):589-607.
- Valdini, Melody E. 2019. *The Inclusion Calculation: Why Men Appropriate Women's Representation*. Oxford, UK: Oxford University Press.
- Valgarðsson, Viktor Orri, Nick Clarke, Will Jennings and Gerry Stoker. 2021. "The Good Politician and Political Trust: An Authenticity Gap in British Politics?" *Political Studies* 69(4):858-880.
- Vivyan, Nick and Markus Wagner. 2015. "What Do Voters Want from their Local MP?" *Political Quarterly* 86(1):33-40.
- Wahman, Michael, Nikolaos Frantzeskakis and Tevfik Murat Yildirim. 2021. "From Thin to Thick Representation: How a Female President Shapes Female Parliamentary Behavior." *American Political Science Review* 115(2):360-378.
- Wang, Vibeke. 2014. "Tracing Gender Differences in Parliamentary Debates: A Growth Curve Analysis of Ugandan MPs' Activity Levels in Plenary Sessions, 1998-2008." *Representation* 50(3):365-377.
- Wängnerud, Lena. 2000. "Testing the Politics of Presence: Women's Representation in the Swedish Riksdag." *Scandinavian Political Studies* 23(1):67-91.
- Wängnerud, Lena. 2009. "Women in Parliaments: Descriptive and Substantive Representation." *Annual Review of Political Science* 12:51-69.
- Ward, Orlanda. 2016. "Intersectionality and Press Coverage of Political Campaigns: Representations of Black, Asian, and Minority Ethnic Female Candidates at the U.K. 2010 General Election." *International Journal of Press/Politics* 22(1):43-66.
- Weeks, Ana Catalano. 2019. "Quotas and Party Priorities: Direct and Indirect Effects of Quota Laws." *Political Research Quarterly* 72(4):849-862.
- Weeks, Ana Catalano. 2022. "The Political Consequences of the Mental Load." *Working Paper* pp. 1-29.
- Weeks, Ana Catalano, Bonnie M. Meguid, Miki Caul Kittilson and Hilde Coffe. 2022. "When Do Männerparteien Elect Women? Radical Right Populist Parties and Strategic Descriptive Representation." *American Political Science Review* (Forthcoming):1-18.

- Weeks, Ana Catalano and Lisa Baldez. 2015. "Quotas and Qualifications: The Impact of Gender Quota Laws on the Qualifications of Legislators in the Italian Parliament." *European Political Science Review* 7(1):119–144.
- Weinberg, James. 2021a. "Emotional Labour and Occupational Wellbeing in Political Office." *British Journal of Politics and International Relations* 23(3):430–450.
- Weinberg, James. 2021b. "Who Wants To Be a Politician? Basic Human Values and Candidate Emergence in the United Kingdom." *British Journal of Political Science* 51(4):1565–1581.
- Welch, Susan. 1977. "Women as Political Animals? A Test of Some Explanations for Male-Female Political Participation Differences." *American Journal of Political Science* 21(4):711.
- Widmann, Tobias and Maximilian Wich. 2022. "Creating and Comparing Dictionary, Word Embedding, and Transformer-Based Models to Measure Discrete Emotional Appeals in German Political Text." *Political Analysis* (Forthcoming):1–16.
- Wilde, Annett and Amanda B. Diekman. 2005. "Cross-Cultural Similarities and Differences in Dynamic Stereotypes: A Comparison Between Germany and the United States." *Psychology of Women Quarterly* 29(2):188–196.
- Winter, Nicholas J.G. 2010. "Masculine Republicans and Feminine Democrats: Gender and Americans' Explicit and Implicit Images of the Political Parties." *Political Behavior* 32(4):587–618.
- Wojcik, Stefan and Shawna Mullenax. 2017. "Men Idle, Women Network: How Networks Help Female Legislators Succeed." *Legislative Studies Quarterly* 42(4):579–610.
- Wolkenstein, Fabio and Christopher Wratil. 2021. "Multidimensional Representation." *American Journal of Political Science* 65:862–876.
- Yildirim, T. Murat, Gülnur Kocapınar and Yuksel Alper Ecevit. 2021. "Status Incongruity and Backlash against Female Legislators: How Legislative Speechmaking Benefits Men, but Harms Women." *Political Research Quarterly* 74(1):35–45.
- Yildirim, Tevfik Murat. 2021. "Rethinking Women's Interests: An Inductive and Intersectional Approach to Defining Women's Policy Priorities." *British Journal of Political Science* (Forthcoming):1–18.
- Yu, Bei. 2013. "Language and Gender in Congressional Speech." *Literary and Linguistic Computing* 29(1):118–132.

Zamani, Hamed and W. Bruce Croft. 2016. "Embedding-based Query Language Models." *Proceedings of the 2016 ACM international conference on the theory of informational retrieval* pp. 147–156.

Ziv, Jacob and Abraham Lempel. 1977. "A Universal Algorithm for Sequential Data Compression." *IEEE Transactions on Information Theory* 23(3):337–343.