Legislative Role Models: Female Ministers, Participation, and Influence in the UK House of Commons

Jack Blumenau

London School of Economics

Abstract

When women are promoted to high office, do they serve as role models to other women in politics? I evaluate a female role-model hypothesis by examining parliamentary debates in the UK House of Commons. In the context of a difference-in-differences design which exploits over-time variation in the gender of cabinet ministers, I demonstrate that appointing a female minister increases the participation of other female MPs in relevant debates by approximately one third, compared to when the minister is male. Further, using a measure of debate influence, based on the degree to which the words used by one legislator are subsequently adopted by other members, I show that female ministers also increase the influence of female backbenchers. To explore the mechanisms behind these results, I introduce a new metric of ministerial responsiveness and show that female ministers are significantly more responsive to the speeches of female backbenchers than are male ministers.

Word count: 9972

> Many years ago I worked in the House of Commons for a woman that I admired very much called Barbara Castle. She was my role model because I felt, well, if Barbara can do it then I can do it. (Boothroyd, 2013)

Baroness Boothroyd, Former Speaker of the House of Commons.

When women are promoted to high political office, do they serve as role models to other women in politics? The factors that determine the appointment of women to political leadership roles have been the subject of increasing study in recent years (Heath, Schwindt-Bayer and Taylor-Robinson, 2005; Escobar-Lemmon and Taylor-Robinson, 2008; Krook and O'Brien, 2012; O'Brien, 2015), but we know considerably less about the implications of these appointments for the behaviour of other politicians. This is surprising as the idea that successful or prominent women can act as 'role models' to other women is well established in electoral politics (Wolbrecht and Campbell, 2007; Beaman et al., 2008; Gilardi, 2015), education (Nixon and Robinson, 1999; Rask and Bailey, 2002; Bettinger and Long, 2005; Dee, 2007; Brajer and Gill, 2010; Beaman et al., 2012) and business (Wang and Kelan, 2013; Bertrand et al., 2014). As the epigraph suggests, the potential for female leaders to act as role models for other women in the legislature is acknowledged by female politicians themselves, but this idea has not previously been subjected to systematic empirical scrutiny.

This paper evaluates a female role-model hypothesis in the context of the UK House of Commons. I focus on one particularly visible leadership role in the legislature – cabinet ministers in the UK government – and study the effects of the appointment of female cabinet ministers on the participation and influence of other female members of parliament (MPs) in plenary debates. When speaking is a mechanism for collective decision-making, inequalities in participation and influence may reflect deeper inequalities between groups (Karpowitz, Mendelberg and Shaker, 2012). Furthermore, the ways in which individuals interact in group discussions can provide important insights into relative distributions of power, particularly with regard to gender (Karakowsky, McBey and Miller, 2004). Understanding the conditions under which female legislators participate and hold influence in political debates is therefore important for evaluating the representation of women's interests in politics more broadly.

I argue that the appointment of a female cabinet minister may affect other female MPs' willingness to participate in debate, and the influence they enjoy as a result of their participation. When women are appointed to high-office, this sends a signal to other female legislators about the benefits to be gained from participating in policymaking, and may help to break down historically constructed stereotypes concerning the appropriateness of female political rule. Furthermore, female ministers' debating styles are likely to be more conducive to, and encouraging of, the participation of other female MPs in parliamentary debate. In the UK, cabinet ministers are highly visible actors in the policy process and play a central role in the parliamentary debates that relate to their ministries, speaking frequently to answer questions and to propose legislation for consideration. Consequently, I expect the appointment of a female minister to lead to higher levels of female participation and influence in debates that are presided over by the new minister.

Causal identification of role-model effects, however, presents an empirical challenge. Ministries to which women are appointed differ in several ways from ministries presided over by men. In particular, the factors that drive the appointment of female ministers to certain ministries are correlated with the probability that women participate in policy debates pertaining to those ministries. For example, women are more likely to be appointed to traditionally "feminine" cabinet posts (Escobar-Lemmon and Taylor-Robinson, 2008; Krook and O'Brien, 2012), and are also disproportionately likely to contribute to debates that deal with traditionally "feminine" policy areas (Taylor-Robinson and Heath, 2003; Bird, 2005; Catalano, 2009). Because of this, simple estimates of the relationship between cabinet minister gender and female debate participation will be upwardly biased.

The strategy here makes progress by exploiting within-ministry variation in the gender of the cabinet minister over time. By assigning each debate to a specific ministry, I compare the level of female debate participation in a ministry before and after a switch in the gender of the minister, and compare this difference to changes in female participation in other ministries where the gender of the minister remains constant. This approach – which is equivalent to a difference-in-differences design in a multi-period setting – allows me to rule

out any omitted variable bias that could be attributed to any fixed tendencies of women to engage with the work of particular ministries and not others. Using this design-based framework to analyse over half a million Commons' speeches between 1997 and 2017, I demonstrate that appointing a female minister increases the *participation* of women MPs in relevant debates by approximately one third over the level of female participation under male ministers.

These results are robust to a number of alternative specifications and provide strong evidence of a role-model effect. However, if women speak more after the appointment of a female minister, but the issues that they raise are ignored by other parliamentarians, then the substantive importance of this effect may be questioned. I address this issue by building on new techniques for identifying important speakers in political debate (Erkan and Radev, 2004; Fader et al., 2007), which I use to examine the *influence* of female MPs. I model the speeches of a parliamentary debate as a directed graphical network, and assess the linguistic centrality of members' speeches within a debate-graph in order to infer how influential each MP is in each debate. The strategy is based on the assumption that the more that an MP's language is adopted by other MPs in subsequent speeches, the more influential is the MP. Using this measure I show that in addition to becoming more loquacious, women also become more influential in debate when their female colleagues are elevated to high-office.

While the mechanisms behind these effects are likely to be varied, I provide evidence that female ministers behave systematically differently towards female MPs than do male ministers. I introduce a new quantitative measure of ministerial responsiveness which assumes that a minister is more responsive when the language they use to reply to a backbencher is more similar to the words that the backbencher uses. I use this measure to demonstrate that female ministers are substantially more responsive than their male counterparts to the speeches made by female MPs.

Role models in the legislature

Existing research provides evidence for a female role-model hypothesis outside of the legislative setting. In countries with higher proportions of female representatives, women are

more likely to discuss politics, and to participate in political activities (Wolbrecht and Campbell, 2007). The election of female politicians also affects the educational attainment and career aspirations of adolescent girls (Beaman et al., 2012) and increases the propensity for other women to stand for elections (Beaman et al., 2008; Gilardi, 2015). In education, assignment to same-sex teachers can significantly improve educational achievement (Dee, 2007; Nixon and Robinson, 1999); influence the course choices of students (Rask and Bailey, 2002; Bettinger and Long, 2005); and improve communication between students and teachers (Brajer and Gill, 2010). Additionally, appointing women to corporate boards can increase the number of women occupying other leadership positions within business (Wang and Kelan, 2013; Bertrand et al., 2014).

Why, then, might the appointment of women to positions of legislative power encourage other female legislators to participate in plenary debate? In the electoral realm, women see themselves as less qualified to run for political office than men (Fox and Lawless, 2011), as less informed about political matters (Mendez and Osborn, 2010) and are less likely to run for office (Fox and Lawless, 2004). If women see themselves as less qualified for office, they may also see themselves as less qualified for participating in the policy process. For example, women tend to stand for, and be elected to, predominantly "feminine" cabinet assignments and committee positions (Escobar-Lemmon and Taylor-Robinson, 2008; Krook and O'Brien, 2012) and female parliamentarians in the UK contribute more to policy debates that concern traditionally feminine issues, but less to debates associated with more masculine policy areas (Bird, 2005; Catalano, 2009). Observing the success of one woman in a policy area may send a signal that women in general are qualified to contribute to that domain (Lockwood, 2006), and may therefore encourage further participation of other women in plenary discussion.

Furthermore, female appointments may break down historically constructed stereotypes regarding the appropriateness of female leadership. Women are systematically underrepresented in leadership positions (Heath, Schwindt-Bayer and Taylor-Robinson, 2005; Krook and O'Brien, 2012); are appointed to leadership roles in unfavourable circumstances

(O'Brien, 2015); and tend to control low prestige portfolios (Studlar and Moncrief, 1999; Russell and DeLancey, 2002). The historical marginalisation of women in high-power roles may create entrenched perceptions that certain policy areas, and even politics in general, represent distinctly "male domains" (Sapiro, 1981, 712). Group-based stereotypes are often the source of negative evaluations of women's capabilities as leaders (Eagly and Johnson, 1990) and exposure to role models can help to undermine stereotypic beliefs (Dasgupta and Asgari, 2004). By breaking with historical patterns, the appointment of women to powerful cabinet positions may therefore reverse the impression that women are unsuitable for participation in politics (Mansbridge, 1999).

In addition, theorists argue that the "deeply embedded culture of masculinity" (Lovenduski, 2005, 48) that pervades parliament is contra-indicated to female influence in political discussion. Legislatures are marked by highly gendered conversational dynamics in which male contributions to policymaking are "heard" more than female contributions (Kathlene, 1994; Hawkesworth, 2003) and, in the UK, the declamatory and adversarial style of Westminster debate (Childs, 2004) is seen as particularly antithetical to the participation and influence of women in the policy-process. Experimental evidence suggests that as the proportion of women in a group decreases, women are likely to be interrupted more often by men (Mendelberg, Karpowitz and Oliphant, 2014); discussion will focus less on traditional "women's issues" (Mendelberg, Karpowitz and Goedert, 2014); and each individual women will speak less (Karpowitz, Mendelberg and Shaker, 2012). These studies also find that when groups operate under majority-rule (as opposed to unanimity), women experience less conducive speaking environments. This is particularly relevant when thinking about parliamentary settings, where women normally constitute a small proportion of discussants, and discussions almost always take place under majoritarian decision rules.

If the appointment of a female politician to a position of high-office changes the experiences of other women in parliamentary discussions, we should expect an increase in women's influence and participation in those debates. For example, women tend to be more democratic in their approach to leadership (Eagly and Johnson, 1990) and in the

legislative setting, female committee chairs act more as moderators or facilitators, rather than directors, of committee discussions, speaking less and make fewer interruptions than their male counterparts (Kathlene, 1994). In the UK context, female politicians in the Commons tend to employ a distinct form of language and debating style (Childs, 2000) which is more cooperative, approachable, and practical than that of their male colleagues (Lovenduski and Norris, 1996). In general, female rhetorical styles are less aggressive, more inclusive, and more cooperative than male speech patterns (Karpowitz, Mendelberg and Shaker, 2012, 534). Women also tend to be characterised by facilitative styles of speech, marked by high levels of politeness and responsiveness, while male speech is seen as less facilitative (Hannah and Murachver, 1999, 2007) and these styles are strongly predictive of the speaking time of conversational partners (Thomson, Murachver and Green, 2001).

An important component of these styles is the degree to which an individuals' contributions to discussion are responded to and acknowledged by other participants (Hannah and Murachver, 2007; Asgari, Dasgupta and Cote, 2010). Accordingly, one potential locus of a role-model effect is in the differential *responsiveness* of male and female high-office-holders. Cabinet ministers have institutionally privileged positions that almost always mean that they speak first in debates, and speak more often than backbench MPs. If female cabinet ministers provide higher quality responses to the speeches of female legislators than do their male counterparts, the status of women in legislative debate is likely to increase when a woman is appointed. Such an increase in status is likely to be concomitant with increases in the degree of influence that women enjoy, and thus their willingness to participate in plenary debate.

Overall, there are numerous reasons why the appointment of a female legislative leader might affect the incentives of female MPs to participate in legislative debate. The central implication that I test in the next section is that when a female MP is appointed to lead a ministry previously held by a man, other female MPs will be more likely to participate in debates that pertain to that ministry than they would have been previously. In addition, as participation and influence in debate are closely related concepts (Kathlene, 1994, 573),

and inequalities in participation between men and women are also associated with inequalities in influence (Karpowitz, Mendelberg and Shaker, 2012), we might also expect that the influence of women in political discussion is also likely to increase when debates are presided over by a high-ranking women. In the subsequent section I investigate whether the appointment of a female minister is also associated with an increase in the level of influence that female MPs experience in plenary debate.

Data, sample, and methodology

I study all House of Commons floor debates between May 1997 and February 2017. The full sample contains 53,397 debates, comprising over a million individual speeches.[1] In this section, I am interested in comparing the volume of speeches delivered by women in debates pertaining to ministry m at time t when the minister for ministry m is female to the counterfactual in which the minister is male.

The key independent variable is the gender of the minister responsible for a ministry at a given point in time. The dummy variable, $FemaleMinister_{mt}$, is equal to one when the minister responsible for a given ministry m in time t is a woman, and zero otherwise. Figure S2 in the appendix shows the variation in this variable over time for all 33 ministries included in the sample.[2] During the study period, there are several ministries for which the responsible minister is never a woman – including the Defence ministry and the position of Chancellor of the Exchequer – but no ministry is always controlled by a female minister. While I include all ministries in the empirical analysis, identification of the role-model effect relies only on those ministries which see a change in the gender of the cabinet minister over time.

The analysis requires that each debate is mapped to an individual ministry. To assign debates to ministries, I note whether a current cabinet minister speaks in a given debate,

[1] This information comes from `theyworkforyou.com`, a public website that catalogues all speeches made by UK MPs.

[2] I define a cabinet minister as any individual who is paid a government salary and regularly attends cabinet meetings. The policy responsibilities for each ministry change occasionally over time and I therefore treat each change in the name of the ministry as a new ministry.

and assign the debate to the ministry for which that cabinet minister is responsible. Where more than one cabinet minister speaks, I assign the debate to the ministry of the most frequently appearing cabinet minister. As ministers speak regularly in the debates for which they are responsible, this serves as an efficient way of categorising the debates. Some debates do not contain speeches from any cabinet ministers,[3] and the final sample for analysis contains 14,388 debates consisting of approximately 530,000 speeches.

The outcome variable is the proportion of words spoken by female legislators in debates pertaining to ministry m in month t:

$$PropWordsWomen_{mt} = \frac{\#\text{ words spoken by women}_{mt}}{\#\text{ words spoken by men and women}_{mt}} \quad (1)$$

I exclude speeches made by the ministers themselves, ensuring that the figures are not artificially inflated by female ministers speaking more after they are appointed. I also remove speeches made by the Speaker of the House of Commons, which are almost exclusively procedural.[4]

Systematic differences between ministries almost certainly affect the degree to which female legislators choose to participate in legislative debate. For example, previous research shows that women are significantly more likely to participate in legislative debates that relate to areas of traditional concern to women, including health care and children and family issues (Catalano, 2009; Pearson and Dancey, 2011). Figure S1 in the appendix suggests that unobserved ministry characteristics such as these are clearly influential in the data here. Women speak significantly more in ministries such as 'Education' and 'Energy and Climate Change', and significantly less in debates pertaining to the 'Defence', 'Foreign', and 'Justice' ministries. If female ministers are disproportionately appointed to ministries in which the rate of female participation is already high, then naive comparisons between debates presided over by female and male ministers are likely to be upwardly biased.

To overcome these problems, I estimate fixed-effects regressions of the following form:

$$PropWordsWomen_{dmt} = \beta_1 * FemaleMinister_{mt} + \lambda_m + \delta_t + \epsilon_{dmt} \quad (2)$$

[3]These are mostly procedural debates.

[4]In appendix section B, I consider several alternative specifications of the dependent variable. The choice between these various measures is inconsequential.

where $PropWordsWomen_{dmt}$ is defined in equation 1, λ_m is a ministry fixed-effect that washes out any omitted variable bias from unobserved ministry characteristics that are fixed over time (such as the degree to which a ministry deals with policy that is traditionally of greater concern to women), δ_t is a year-month fixed-effect to control for common shocks across ministries in a given month, and ϵ_{dmt} is an idiosyncratic error term. β_1 is the coefficient of interest, and captures the reduced-form causal effect of the appointment of a female minister on the participation of women in debates for those ministries that experienced a change in minister gender over time.

This fixed-effect design is equivalent to a multi-period 'difference-in-differences' in the style of Angrist and Pischke (2009, 234). β_1 identifies the effect of switching from a male to female minister based on the within-ministry variation of the outcome variable among those ministries that see changes in the gender of the minister over time. By accounting for fixed characteristics of ministries that might predict both female debate participation and the appointment of a female minister, the model compares changes in female debate participation in ministries that experience a switch in minister gender to ministries where the gender of the minister remains constant, while differencing out the general trends across ministries in a given month.

Identification of the causal effect relies on changes in minister gender being exogenous to the level of female debate participation, conditional on time and ministry fixed-effects. The key identifying assumption is that treated ministries would have followed the same trend as non-treated ministries in the absence of treatment. I relax this assumption by estimating further models which include ministry-specific linear (λ_{m1}) and quadratic (λ_{m2}) time trends:

$$\begin{aligned} PropWordsWomen_{dmt} &= \beta_1 * FemaleMinister_{mt} + \lambda_{m0} + \delta_t \\ &\quad + \lambda_{m1} t + \lambda_{m2} t^2 + \epsilon_{dmt} \end{aligned} \tag{3}$$

where t is a time variable. Furthermore, in contrast to the typical multi-period 'difference-in-difference' model, in this setting the treatment (the presence of a female minister) switches on and off over time. That is, once appointed, a female minister might also leave

office, and ministries often see multiple female ministers (appointed at different times) over the study period. To account for the possibility that differential *local* trends within ministries might confound the causal effect, I also estimate generalised additive models (GAM) which include non-parametric ministry-specific time trends:

$$\begin{aligned} PropWordsWomen_{dmt} &= \beta_1 * FemaleMinister_{mt} + \lambda_{m0} + \delta_t \\ &\quad + \lambda_{m1} f(t) + \epsilon_{dmt} \end{aligned} \tag{4}$$

These models represent extremely conservative specifications, as the addition of the ministry-specific trends means that all unobserved and smoothly varying confounding differences are removed from the estimate of β_1, and that only sharp changes to the trend in the outcome variable that occur at the same time as the change in minister gender contribute to this estimate. As none of the substantive or statistical results change noticeably when this crucial identifying assumption is relaxed, this lends significant support to the empirical design. I provide further evidence for the validity of the identification assumption by estimating a dynamic panel model, which estimates the treatment effect in the time periods before and after the actual change in minister gender. This allows me to test whether the treated and non-treated ministries experienced systematically different levels of female debate participation in the periods leading up to the treatment. The results from the dynamic model suggest that this is not the case, and thus reduce concerns that the effect I observe is a feature of some factor other than the appointment of a female minister. Finally, as there are only 33 ministries in the data, I follow standard practice in the literature and construct bootstrapped clustered standard errors at the ministry level (Cameron and Miller, 2015).[5]

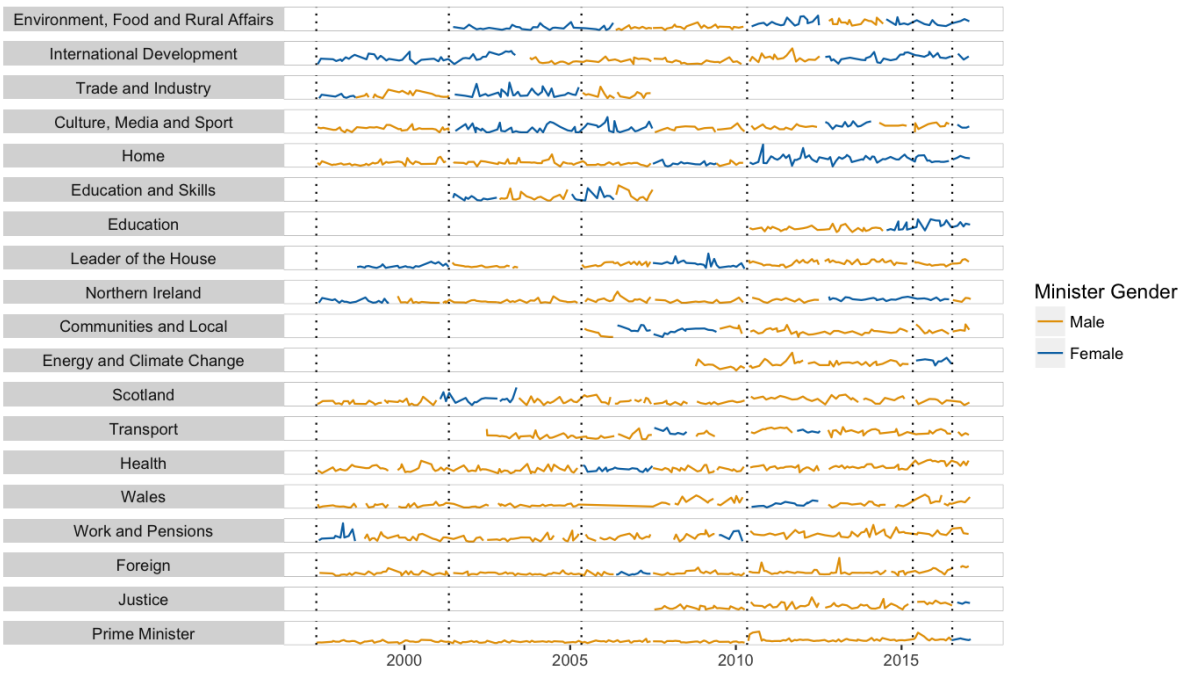
[5]I bootstrap 1000 times, resampling ministries from the full data with replacement, and estimating equations 2 and 3 at each iteration. Because the GAM model is computationally very burdensome, I do not bootstrap this model. Nevertheless, the GAM model provides point estimates which are very similar to the other models.

Female ministers and debate participation

Before turning to the main results, I present a simple graphical analysis. Figure 1 shows the evolution of the dependent variable over time in the ministries that experienced a change in the gender of the minister. The y-axis in the figure gives the proportion of words spoken by female MPs in each month, and the x-axis gives the date. Blue line segments represent periods in which the presiding minister is female, and orange segments represent male ministers. The plot provides clear evidence of a role-model effect whilst also revealing heterogeneity across ministries. In many cases, the appointment of a female minister is accompanied by an increase in the proportion of words spoken by other female MPs. The effect appears to be particularly pronounced in the 'Trade and Industry', 'Home', 'Culture, Media and Sport', and 'International Development' ministries. By contrast, there is less evidence of an effect in other ministries, though in no cases does the appointment of a female minister appear to lead to a *decrease* in the proportion of words spoken by other female MPs.

Table 1 presents the results of the regression analyses. Model 1 presents the naive estimate of the effect of a female minister, without controlling for ministry or year-month fixed-effects. Models 2 and 3 introduce these fixed-effects separately, and model 4 presents the results of the 'difference-in-differences' model which includes both fixed-effects. The coefficient of the main variable of interest, $FemaleMinister$, is positive and significant in all four of these models, but it decreases noticeably when accounting for ministry. This suggests that female ministers are indeed appointed to lead ministries where the level of debate participation of other female MPs is already high. Nevertheless, the effect remains significant in the specification in model 4, implying that the appointment of a female minister leads to an increase in debate participation of other female MPs in the House of Commons. The size of the effect is also substantial. Based on model 4, the appointment of a female minister increases the proportion of words used by women by 4.4 percentage points of total words. This corresponds to an increase of approximately 26% [11%, 41%] over the average speech rate of women in debates under male ministers.

Figure 1: Proportion of words spoken by women in treated ministries, over time



NOTE: The plot shows the proportion of words spoken by women in each calendar month, in each ministry that experienced a change in the gender of the presiding minister.

Table 1: Effect of appointing a female minister on female debate participation

	PropWordsWomen						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female minister	0.060***	0.054***	0.056***	0.044***	0.040***	0.041***	0.033***
	(0.017)	(0.016)	(0.016)	(0.013)	(0.013)	(0.011)	(0.008)
Constant	0.169***	0.098***	0.109*	0.055	0.119	0.159	0.038
	(0.011)	(0.025)	(0.066)	(0.034)	(2.015)	(552.866)	(0.377)
Month FEs	×	✓	×	✓	✓	✓	✓
Ministry FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Effect Size %	36	32	33	26	23	24	19
95% CI	[16,56]	[13,50]	[14,52]	[11,41]	[9,38]	[12,37]	[10,29]
Observations	3,057	3,057	3,057	3,057	3,057	3,057	3,057
R^2	0.034	0.240	0.192	0.366	0.395	0.416	
Adjusted R^2	0.034	0.184	0.184	0.312	0.336	0.352	0.404

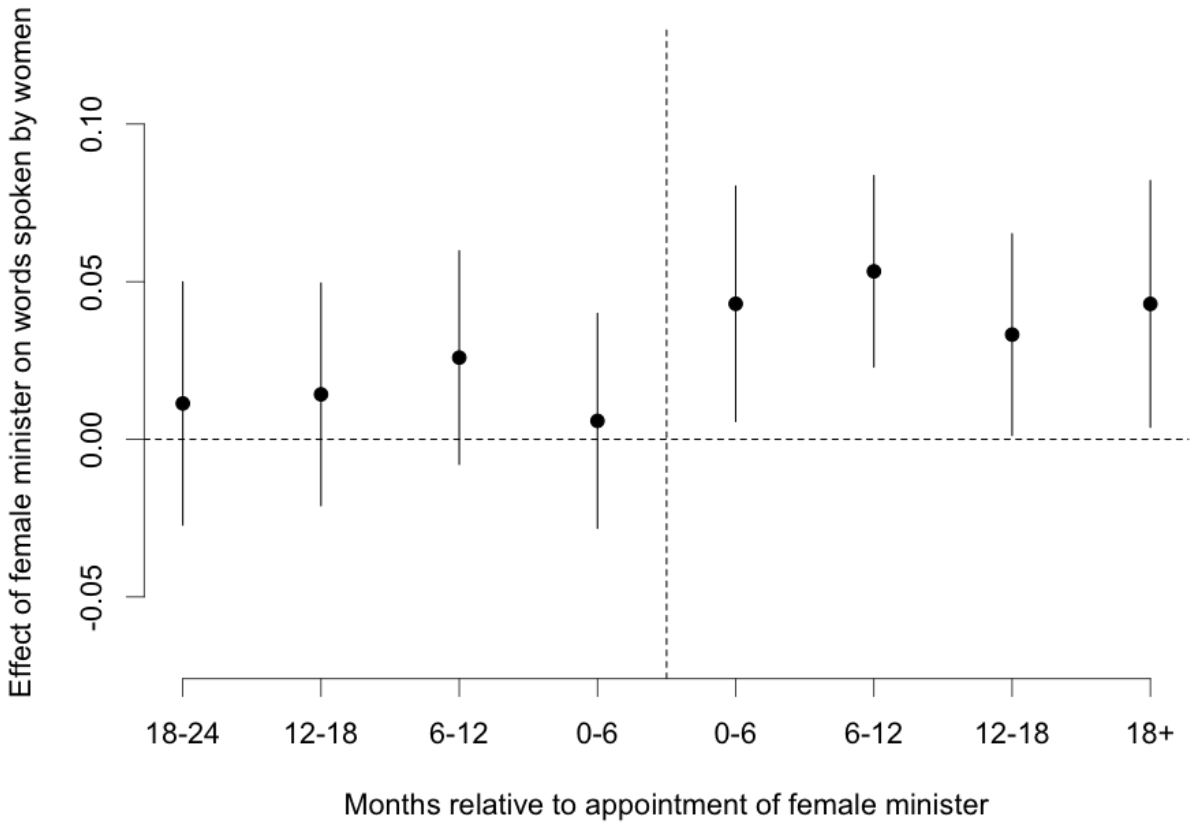
NOTE: Models 1-6 represent OLS fixed-effect regressions for the period 1997-2017. Regression coefficients are shown with bootstrapped cluster-robust standard errors (clustered by ministry) shown in parentheses. The "Effect Size" row indicates the percentage increase in female participation relative to the average female participation rate under male ministers. *p<0.1; **p<0.05; ***p<0.01

The main effect is also robust to a number of alternative specifications. Models 5, 6 and 7 introduce linear, quadratic and non-parametric ministry-specific time trends. As stated above, the inclusion of these trends relax the crucial identifying assumption that treated and non-treated ministries would have followed parallel trends in the absence of treatment. The table shows that the effects are highly robust, and the effect size is consistent: based on model 6, the appointment of a female minister increases the proportion of words spoken by other female MPs by between 12% and 37%. That the inclusion of ministry-specific time trends changes the estimates so little is encouraging, as it rules out the possibility that the documented effect is driven by either global or local trends in unobserved confounding variables.

As a final robustness check, figure 2 plots the coefficients estimates from the dynamic panel model. Here I code a binary indicator for the first 6 months of the treatment period in a given ministry, and then add four leads and three lags of this indicator in addition to the full set of fixed-effects and linear and quadratic time-trends. The final lagged variable captures all treated periods from eighteen months until the end of the treatment period for a given ministry. The coefficients represent the estimated difference in the outcome between treated and untreated ministries in the periods before and after the treatment occurs. The results strongly support the identifying assumption, as I find no significant 'placebo' effects in the two years prior to the change in minister gender. This strengthens the plausibility of the design, as it suggests that there are no unobserved variables leading to differential trends in the outcome between the treatment and control ministries prior to the appointment of a female minister. Furthermore, there is a large and significant difference between treatment and control ministries in all periods after a female minister is appointed.

Overall, the results presented here provide strong support for a female role-model hypothesis: when a female minister is appointed, other women speak approximately one third more in debates pertaining to that ministry than when the responsible minister is male.

Figure 2: Dynamic panel model estimates



NOTE: The plot presents estimates of switching from a male to a female minister before and after the actual change occurred. The vertical dashed line indicates the timing of the change, and the points indicate (at six month intervals) the difference between treated and untreated ministries at the given time point.

Female ministers and influence

The findings in the previous section indicate that female backbenchers are more likely to participate in the policymaking process when the responsible minister is a woman. However, changes in relative levels of participation tells us little about how these debate contributions are received by others in the House. If women are speaking at increasing rates, but the issues and concerns that they raise are largely ignored by other parliamentarians, then the substantive importance of the documented role-model effect may be limited. In this section, I use the texts of the parliamentary speeches to provide evidence that female backbenchers do not only speak more after the appointment of a female minister, but also that they play a more influential role in political debate than under male ministers.

Which features of text might identify 'influential' speakers? Rather than directly analysing the content of each speech, I study the links *between* speeches in order to assess the influence of any particular speaker in a debate. I consider an MP to be influential when the issues and concerns she raises in her speeches are adopted and discussed by other members in subsequent speeches. Having other people pick up on your framing of an issue is a way of controlling how the debate proceeds: it means other people are taking up your perspective, whether or not they agree with it. Influential MPs are therefore literally 'shaping the debate'. The intuition behind the measurement strategy is to identify distinctive language that first appears in the statement of one MP, but then gets used subseqeuntly by later MPs.

I build upon methods for detecting influence in text corpora, where the corpus of texts is treated as a graphical network, and influential documents are identified by their lexical centrality in the network (Erkan and Radev, 2004; Mihalcea, 2004; Fader et al., 2007). Debates can be viewed as clusters of documents that are related to each other in the language that they use. Some speeches within a debate will share similar language, while others will share less information with one another. The crucial intuition behind the model employed here is that those speeches that are similar to many other speeches within the same debate are more central to the topic, and speakers of central speeches are more

influential in a legislative debate.

I assess the influence of a speech, i, by calculating how many 'references' i receives from other speeches within the debate.[6] One speech, j, can be understood to 'reference' another, i, when it occurs after i in the debate and when it comprises language which is sufficiently similar to that used by i. By using similar language to i, j is implicitly indicating that i is relevant and important for the discussion at hand. There are myriad reasons why one speech may use similar language to another (direct quotation; expression of criticism; statement of support) but the goal here is not to assess the substantive meaning of each link. Rather, I assume that a speech that shares language with many other speeches is being *collectively referenced* and thus can be considered an important and influential speech within the debate. A simple way of assessing influence would therefore be to simply count the number of references each speech receives. I consider not only the number of references a given speech receives, but also incorporate information about the influence of the referencing speeches. Thus, the more references i receives, and the higher the influence of the referencing speeches, the higher is the influence i within the debate.

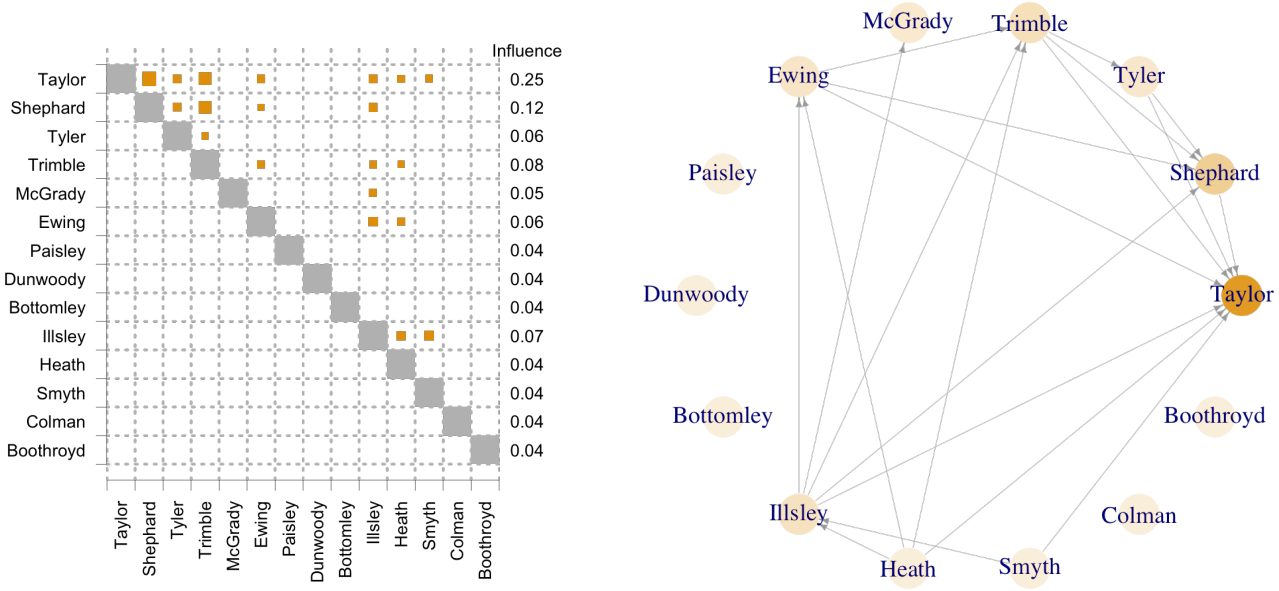
The basic steps of the measurement procedure are as follows.[7] First, I construct debate-specific similarity matrices which measure the cosine-similarity between all speeches in each debate, where speeches are represented as term-frequency-inverse-document-frequency (tf-idf) vectors. Second, these matrices are converted into directed graphs (again, one for each debate) where the nodes represent speeches, and edges are placed between nodes for the speech-pairs whose cosine-similarity is greater than some threshold value.[8] The edges are then weighted by the similarity scores. Third, I analyse these matrices using an iterative ranking algorithm (Page et al., 1999) to calculate a vector of centrality scores, P, which correspond to the influence of each *speech* in each debate. Finally, the influence score of an *MP* in a given debate is the sum of the influence scores for the speeches given by that MP in that debate.

[6] I use the term 'reference' as a heuristic for describing the tendency for the vocabulary used in the speech of one member to be subsequently used in the speeches of other members.

[7] Full details can be found in appendix section C.

[8] In line with Fader et al. (2007) I set this threshold to 0.25.

Figure 3: Example of 'influence' in a debate



NOTE: The left panel shows the 'reference' patterns for an example debate. The right panel depicts the similarity matrix as a network graph.

In figure 3, the left-hand plot depicts the similarity matrix for an example debate with 14 separate speeches. Speakers are sorted according to the order in which they participated in the debate, such that Taylor is the first speaker and Boothroyd is the last speaker. As I only allow one speech to reference another when it occurs later in the debate than the speech it is referencing, the bottom triangle of the matrix is empty. I also exclude the possibility that a speaker can reference herself (grey shaded boxes). The orange-shaded squares indicate the cosine similarity between two speeches, and are scaled such that when the similarity between a pair of speeches is 1 (i.e. when the tf-idf vectors are identical) the orange square will fill the dashed box that contains it. The empty elements of the upper triangle correspond to speech pairs where the similarity between the speeches is lower than the minimum threshold. The left panel shows, for example, that Taylor's speech is referenced by many subsequent speeches, while Tyler's speech is referenced only by Trimble. The right margin of the plot gives the vector of influence scores for this debate. Taylor's speech has an influence score of 0.25 and Tyler's speech has an influence score of just 0.06.

The right-hand panel depicts the same similarity matrix as a directed network graph, with speeches as nodes and edges as the 'references' flowing from one speech to another. Taylor's speech is referenced by many other members, while Bottomley, Colman, Dunwoody, and Paisley make speeches which are not sufficiently similar to any subsequent speeches and therefore these speakers receive no references.

Validation is essential for text-based measures of political concepts (Grimmer and Stewart, 2013) and in appendix table S5 I show that cabinet ministers (who have agenda-setting privileges in debate) are on average 5 times more influential than backbench MPs, while the Speaker of the House (whose speeches are almost exclusively procedural) is less than half as influential as backbenchers. In addition, in appendix figure S4 I show that the influence scores estimated from the procedure above correlate strongly and positively with the number of times that an MP is directly mentioned by other MPs in debate. These comparisons provide reassuring evidence regarding the face validity of the measure of influence described above.

One potential concern is that this measure of influence is simply proxying for speech *length*. Additionally, because references only flow from later speeches to earlier ones, there may be a relationship between influence and the *position* that a speech occurs in a debate. In figure S5 in the appendix I investigate the relationship between influence and both speech length and debate position across all debates in the sample. As the figures show, there is a very weak relationship between length and influence (the average correlation across all debates is -0.05), and although there is a stronger negative association between influence and debate position, the influence measure is clearly picking up information above and beyond simple debate ordering (the average correlation is -0.45).

With this measure in hand, I now analyse the effect of the appointment of a female cabinet minister on the influence of female MPs. In contrast to the previous analysis here I concentrate on changes at the individual – rather than debate – level. As before, I exclude all speeches made by cabinet ministers and by the Speaker of the House and estimate

models of the form:

$$\begin{aligned} influence_{imt} &= \beta_1 * FemaleMP_i + \beta_2 * FemaleMinister_{mt} + \\ &\quad \beta_3 * (FemaleMP_i * FemaleMinister_{mt}) + \\ &\quad \lambda_{m0} + \lambda_{m1} t + \lambda_{m2} t^2 + \delta_t + \epsilon_{s(imt)} \end{aligned} \tag{5}$$

where $influence_{imt}$ represents the influence of member i in debates pertaining to ministry m at time t. $FemaleMP_i$ is a binary variable, equal to one when the MP is female, and zero otherwise. $FemaleMinister_{mt}$ is equal to one when the minister responsible for ministry m and time t is female, and zero otherwise. β_1 captures the average difference in influence between male and female MPs when the minister is male. β_2 represents the marginal effect of a female minister on the influence of male MPs, and the equivalent effect for female MPs – and the main quantity of interest – is given by $\beta_2 + \beta_3$. The role-model hypothesis implies that $\beta_2 + \beta_3 > 0$, i.e. that female MPs' influence increases after the appointment of a woman minister. As previously, in addition to ministry and time fixed-effects (λ_{m0} and δt, respectively), I also relax the common trend assumption with the addition of ministry-specific linear (λ_{m1}), quadratic (λ_{m2}), and non-parametric time trends. Errors are again clustered at the ministry level. Table 2 presents the results.

Model 1 gives the results of a naive specification without controlling for ministry or time fixed-effects, and indicates that while the appointment of a female minister has no effect on the influence of *male* MPs in parliamentary debates, *female* MPs' influence does increases when a female minister is appointed. The introduction of ministry and time fixed-effects in models 2, 3 and 4, does not change the estimate dramatically: the appointment of a female minister is significantly related to an increase in the influence of female MPs in parliamentary debate, but has no effect on the influence of male MPs. Models 5, 6 and 7 include ministry-specific linear, quadratic and non-parametric time trends. As before, that the effect does not disappear once controlling for these trends gives additional support to the design-based identification strategy.

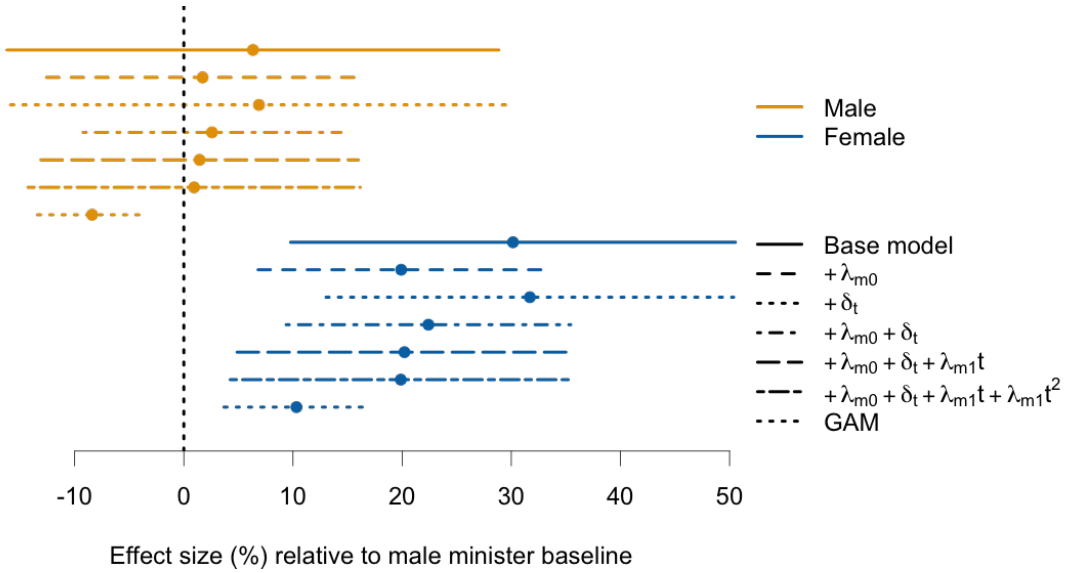
Figure 4 shows the percentage change in influence for male (orange lines) and female (blue lines) MPs after the appointment of a female minister, relative to a baseline where

Table 2: Effect of appointing a female minister on MPs' debate influence

	Influence						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	0.047***	0.047***	0.055***	0.033*	0.051***	0.045***	35.838
	(0.004)	(0.0002)	(0.015)	(0.018)	(0.014)	(0.016)	(164.412)
Female minister	0.003	0.001	0.003	0.001	0.001	0.0004	−0.004***
	(0.005)	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)	(0.001)
Female MP	−0.001	−0.001	0.001	0.001	0.001	0.001	0.001
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
Interaction	0.011***	0.008***	0.011***	0.009***	0.008***	0.009***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)
Ministry FEs	×	✓	×	✓	✓	✓	✓
Month FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Observations	174,419	174,419	174,419	174,419	174,419	174,419	174,419
R^2	0.001	0.042	0.015	0.054	0.060	0.059	
Adjusted R^2	0.001	0.042	0.014	0.052	0.058	0.058	0.075

NOTE: Models 1-6 present OLS fixed-effect regressions for the period 1997-2017, model 7 presents results from the GAM. Regression coefficients are shown with robust standard errors (clustered by ministry) shown in parentheses. *p<0.1; **p<0.05; ***p<0.01

Figure 4: Marginal effect of female minister on influence



NOTE: The plot shows the marginal effect of the appointment of a female cabinet minister on the debate influence of male (orange lines) and female (blue lines) MPs, relative to the average level of influence when the minister is male.

the minister is male. The marginal effect for male MPs is close to zero, varying in sign, and statistically insignificant for all models except for the GAM. For female MPs, the effect is always positive and significant, and the magnitude is non-trivial: based on model 6, female MPs are 20% [4%, 35%] more influential under a female minister than when the minister is male. In sum, the results indicate that, consistent with a role-model effect, the appointment of a female minister leads to an increase not only in the degree to which female MPs participate in plenary debate, but also in the level of influence that female MPs enjoy when debating with their fellow parliamentarians.

MINISTERIAL RESPONSIVENESS

The analysis in the previous sections demonstrated reduced-form evidence of a female role-model effect in the House of Commons, both in terms of participation and influence. How might we account for these effects? The processes underpinning these findings are likely to be many and varied, and isolating the mechanisms behind causal effects is notoriously

difficult with observational data (Imai et al., 2011). However, one explanation for this finding might be that female cabinet ministers behave in a systematically different manner towards female MPs than do male ministers. In particular, female ministers may be more *responsive* to the speeches of female MPs. Such a hypothesis is consistent with findings in the literature on social linguistics: "conversational partners who offer encouragement and are attentive and responsive are more likely to elicit frequent and active participation from speakers in the conversation." (Hannah and Murachver, 1999, 157) If female ministers give other female MPs responses to their speeches that are of higher quality, this is likely to serve as a signal that they are valued colleagues in the House, and may encourage higher levels of participation in future debates. Similarly, the measure of influence outlined in the previous section is based on the idea that speakers are more influential when the language that they use is adopted by other members is subsequent speeches. Accordingly, if female cabinet ministers provide fuller responses to the contributions of female MPs than do male ministers, then this could directly explain the empirical results presented above.

What are the important properties of responsiveness? First, a response is the occurrence of one phenomenon after the occurrence of another phenomenon. Second, a responsive speech is necessarily reactive, and involves engaging with or replying to concepts raised in an original speech. To measure 'responsiveness', I therefore assume that a speech, j, responds to another speech, i, when it occurs directly after i and when it engages with the same thematic content as i. I also consider one speech to be more responsive to another when that speech is longer, on the assumption that longer responses provide more information to the original speaker than shorter ones, and give a greater impression of attentiveness and fullness of reply than shorter responses.[9]

I define a metric which measures how similar two (consecutive) speeches are in terms of the words that they use. Making use of the vector-representation of speeches described

[9]Eggers and Spirling (2014) evaluate ministerial responsiveness by analysing the relative frequency with which ministers speak after backbenchers in debate. By contrast, I focus on the degree to which – conditional on a minister speaking after a backbencher – the language used by a minister is similar to that used by the backbencher.

in equation S4, the responsiveness of speech j to speech i is given by:

$$res_{j \to i} = sim(i, j) * n_j \tag{6}$$

where the first term on the right-hand side of the equation is the cosine-similarity between the two tf-idf vectors, and n_j is the number of words in speech j. When all elements of v_i and v_j are positive, as they are here, the cosine-similarity of two documents is bounded between zero and one. An intuitive interpretation of $res_{j \to i}$ is therefore the (weighted) number of words in speech j that are responding to speech i.[10]

I provide two types of validation for this measure in appendix section F. First, I show that equation 6 captures something distinct from topicality, as comparing pairs of speeches *within the same debate*, those speeches that follow directly after each other are more responsive than speeches that are non-adjacent. Second, I demonstrate that patterns of responsiveness conform with basic intuitions of parliamentary behaviour in the Commons by leveraging the particular structure of minister-backbencher interactions in Question Time debates.

I now turn to the main analysis. To reiterate, if female MPs speak more and become more influential because they receive higher quality responses from female ministers than male ministers, then ministerial speeches subsequent to female speeches should be marked by higher levels of res when the presiding minister is female. I therefore subset the data to those speeches made by backbench MPs which are immediately followed by speeches made by ministers, and estimate models of the following form:

$$\begin{aligned} res_{s(imt)} \quad = \quad & \beta_1 * FemaleMP_i + \beta_2 * FemaleMinister_{mt} + \\ & \beta_3 * (FemaleMP_{mt} * FemaleMinister_{mt}) + \\ & \lambda_{m0} + \delta_t + \lambda_{m1} t + \lambda_{m2} t^2 + \epsilon_{s(imt)} \end{aligned} \tag{7}$$

The unit of analysis in these models is a speech made by a backbencher, which is immediately followed by a speech made by a minister. Thus, $res_{s(imt)}$ is the response *received* by

[10]Note that as i occurs prior to j, it therefore cannot be understood to 'respond' to j. For this reason, $res_{i \to j}$ is not meaningful in our context, and I calculate equation 6 only for sequentially adjacent speeches.

a speech s made by MP i in a debate pertaining to ministry m at month t. $FemaleMP_i$ is a binary variable, equal to one when the MP is female. $FemaleMinister_{mt}$ is equal to one when the minister responsible for ministry m and time t is female. β_1 indicates the difference in responsiveness received by male and female MPs when the minister is male. β_2 captures the effect of the appointment of a female minister on the responses received by male MPs. β_3 therefore captures the interaction between the gender of the MP speaking, and the gender of the minister responding. A positive value for $\beta_2 + \beta_3$ would indicate that the appointment of a female minister leads to an increase in ministerial responsiveness to speeches by female MPs. As before, I include ministry and time fixed-effects and various ministry-specific time trends. Also as before, errors are clustered at the ministry level.

Table 3 presents the results. The interaction effect of interest, β_3, is positive, significant, and sizeable in magnitude across all model specifications. I plot the substantive magnitude of these effects in figure 5, where the baseline is the average responsiveness of male ministers to male and female speeches. Based on the estimates in model 6, the appointment of a female minister increases the responsiveness to female speeches by 21% [8%, 33%] . By contrast, the appointment of a female minister has no consistent effect on the responsiveness to male speeches: across all models in table 3, β_2 is small in magnitude and in many cases statistically indistinguishable from zero. However, when the first speaker is a woman, then the gender of the responding minister matters.

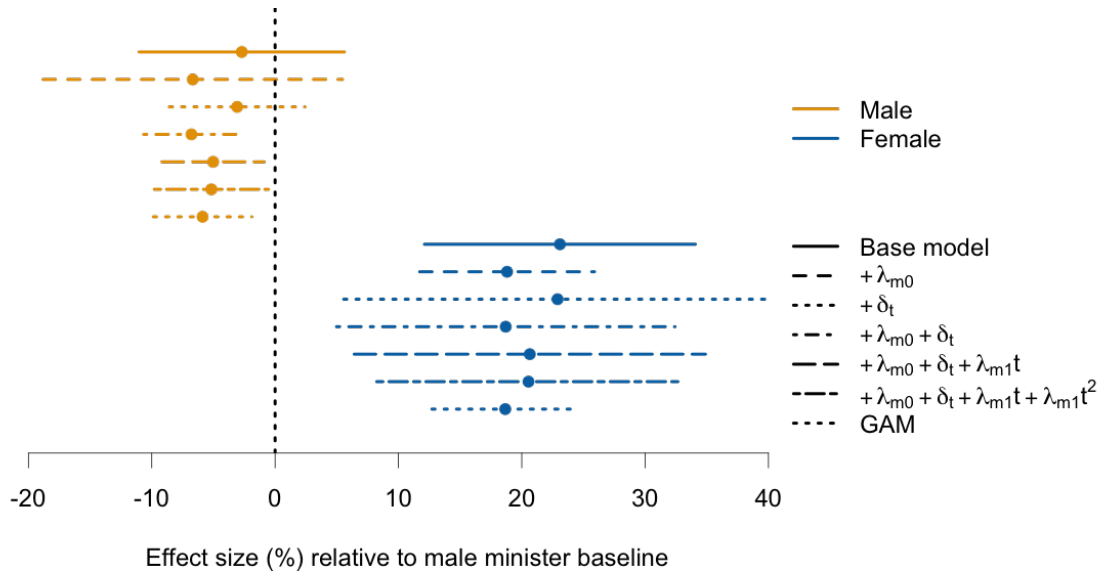
That female MPs receive systematically different responses from male and female ministers helps to explain the increase in influence of female MPs detailed in the previous section. Female MPs become more influential in parliamentary debate (the language that they use in debate is adopted more often in subsequent speeches) after the appointment of a female minister, and this effect is at least partially driven by higher levels of responsiveness of the female minister herself. This may also explain the increase in the participation of female MPs in debate, as higher levels of ministerial responsiveness indicate that the concerns of female MPs are receiving more attention from powerful government figures, and send a signal that the issues that female MPs raise are worthy of governmental concern.

Table 3: Effect of appointing a female minister on the responsiveness to MPs' speeches

	Responsiveness						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	−0.432	−0.210	0.444	0.413	0.462	0.467	0.582**
	(1.595)	(1.609)	(1.561)	(1.585)	(1.586)	(1.587)	(0.288)
Female Minister	−0.626	−1.549	−0.714	−1.574***	−1.166**	−1.199**	−1.365***
	(0.985)	(1.437)	(0.651)	(0.458)	(0.491)	(0.548)	(0.472)
Interaction	5.887***	5.834***	5.931***	5.834***	5.871***	5.879***	5.616***
	(1.794)	(1.736)	(1.802)	(1.762)	(1.796)	(1.757)	(0.645)
Constant	23.211***	22.937***	29.668***	25.695***	27.130***	25.644***	259.757
	(0.700)	(0.164)	(5.335)	(4.838)	(4.514)	(4.553)	(440.085)
Ministry FEs	×	✓	×	✓	✓	✓	✓
Month FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Observations	159,464	159,464	159,464	159,464	159,464	159,464	159,464
R^2	0.001	0.006	0.010	0.014	0.016	0.016	
Adjusted R^2	0.001	0.006	0.009	0.012	0.014	0.014	0.017

NOTE: Models 1-6 present OLS regressions for ministerial responses, and model 7 presents the results of the GAM. Regression coefficients are shown with cluster-robust standard errors in parentheses (clustered on ministry). *p<0.1; **p<0.05; ***p<0.01

Figure 5: Marginal effect of a female minister on responsiveness



NOTE: The plot shows the marginal effect of the appointment of a female minister on the responsiveness to speeches by male (orange lines) and female (blue lines) MPs, relative to the average level of responsiveness when the minister is male.

ALTERNATIVE EXPLANATIONS

I do not claim that the differential level of responsiveness is the only mechanism through which the role-model effect could operate. It is possible, for example, that these results may be driven by the purely symbolic effects of appointing a female minister rather than any behavioural differences. More concerning is the possibility that the reasons for the increase in female participation and influence are completely distinct from the hypothesised role-model effect.

First, if appointing women to visible positions confers a political advantage to the governing party, opposition parties may respond to the appointment of a female government minister by strategically appointing a woman to lead the competing *shadow* ministry (Matland and Studlar, 1996). If this is the case, the documented effects may be due to the fact that institutional rules give both ministers and shadow ministers more time to speak on the House floor than other MPs. I investigate this hypothesis in appendix section G by analysing the relationship between the gender of a *newly appointed shadow minister* and

the gender of the *current government minister* in a given ministry. If opposition parties are behaving strategically, then the probability that a female shadow minister is appointed should be positively associated with the presence of a female government minister. As my regressions show, there is little empirical support for such an argument. In addition, table S9 in the appendix re-runs the main analysis when excluding speeches made by shadow ministers. The results are substantively and statistically very similar to those from the full sample.

Second, ministerial positions come with significant agenda-setting powers, and ministers determine the direction and substance of legislation deriving from their ministries. One possible explanation for the increase in female participation and influence is that female ministers may propose legislation that focuses on topics which are traditionally of greater interest to women, and on which women's contributions are more respected. Such a finding would undermine the notion that female ministers are acting as role models, and suggest instead that female MPs increase their level of participation thanks to a substantive change in the legislative agenda. In section H of the appendix, I examine whether topics that are typically associated with high levels of female participation become more prevalent when a female minister is appointed. I use statistical topic models to estimate which legislative topics are associated with high levels of female participation under male ministers, and then assess the degree to which these topics increase when a female minister takes office. I find no evidence that female ministers are disproportionately introducing legislation that is traditionally associated with high levels of female participation.

Conclusion

In a debate in the House of Commons on International Women's Day in 2003, Joan Ruddock, a Labour Party MP, paid tribute to the minister for Women and the Secretary of State for Trade and Industry, Patricia Hewitt:

> "It is a great pleasure, as it always is, to participate in a debate that is dominated by women. I want to start by congratulating my right honourable

> Friend the Minister for Women, who is a great role model for us – not just in her role as Minister for Women... but by being the head of a substantial Government Department." (Ruddock, 2003)

That female parliamentarians themselves recognise the importance of role models in the legislature makes it all the more surprising that such a premise has been missing from the academic literature. In this paper, I provided evidence for a female role-model effect in the House of Commons by showing that the appointment of a female cabinet minister is associated with an increased propensity for other female MPs to speak in parliamentary debates under the jurisdiction of the new minister. Further, I find that female MPs also become more influential in debates under the purview of female ministers. While there are many mechanisms through which a role-model effect may operate, I show that female ministers respond in a systematically different fashion to the speeches of female MPs than do male ministers.

Legislatures are hierarchical institutions in which some actors have access to positions which confer important powers to the office holder. While the general consequences of these institutional powers are well-studied, there has been less written about the implications of female occupation of such roles for the representation of women. Cabinet posts, committee chairs, and other high-profile legislative offices are normally marked by high levels of visibility and prestige, and make the politicians that hold these posts natural focal points for the public, but also for other members of the legislature. The results here suggest that when women hold high-profile offices, they have significant effects on the behaviour of other legislators, and, crucially, that their appointment can increase the voice of *other* women in the policy process.

Furthermore, a central focus of legislative scholars studying gender issues has concerned the link between descriptive representation – the number of women elected to parliament – and substantive representation – the incorporation of women's interests into policy outcomes (Wängnerud, 2009). While there is growing empirical evidence for this link, the mechanisms that connect increasing numbers of women in parliament to qualitative changes

in political outcomes have not been fully articulated (Beckwith and Cowell-Meyers, 2007). One possibility is that it is not merely the number of women who gain elected office that matters for substantive representation, but also the heights to which those women rise once they have been elected. A limited amount of experimental research indicates that when women take leadership roles, collective decisions tend to reflect distinctly female preferences (Humphreys, Masters and Sandbu, 2006). More research is needed to examine the down-stream effects of female leadership on policy outcomes, but the findings here indicate a possible mechanism through which policy change may occur: female leaders promote increased participation and influence of other women in policymaking. Tracing out a full causal relationship between female leadership and policy outcomes that enhance the substantive representation of women is a difficult empirical task, but the results here suggest, at the very least, that the appointment of women to high-office can have non-negligible effects on the behaviour of other legislators.

Future work should consider the potential for a role-model effect for other disadvantaged groups in different legislative settings. Historically, political elites have disproportionately shared characteristics of the dominant groups in society, and several groups remain significantly underrepresented in the policy process. It would be profitable in the US case, for example, to examine whether the elevation of African-American members to senior positions in the Congressional hierarchy is associated with a concomitant increase in the participation of black legislators in policymaking.

Finally, a growing formal literature examines the consequences of leaders' communication strategies in collective decision making but the empirical literature on communication and leadership has lagged behind (Ahlquist and Levi, 2011). In part, this is due to the difficulty of operationalising reliable measures of spoken communication and establishing credible identification strategies that isolate the effects of leaders in observational settings. This paper makes progress on both fronts. First, the identification strategy I employ suggests that by exploiting variation over time in the identity of political leaders, it is possible to estimate causal effects of leadership on parliamentary outcomes. Second, the measures of

influence and responsiveness introduced here could be profitably applied to other questions of rhetoric and parliamentary leadership. I leave such endeavours for future work.

References

Ahlquist, John S and Margaret Levi. 2011. "Leadership: What it means, what it does, and what we want to know about it." *Annual Review of Political Science* 14:1–24.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Asgari, Shaki, Nilanjana Dasgupta and Nicole Gilbert Cote. 2010. "When Does Contact with Successful Ingroup Members Change Self-Stereotypes?" *Social Psychology* 41(3):203–211.

Beaman, Lori A, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande and Petia Topalova. 2008. Powerful women: does exposure reduce bias? Technical report National Bureau of Economic Research.

Beaman, Lori, Esther Duflo, Rohini Pande and Petia Topalova. 2012. "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India." *science* 335(6068):582–586.

Beckwith, Karen and Kimberly Cowell-Meyers. 2007. "Sheer numbers: Critical representation thresholds and women's political representation." *Perspectives on politics* 5(03):553–565.

Bertrand, Marianne, Sandra E Black, Sissel Jensen and Adriana Lleras-Muney. 2014. Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway. Technical report National Bureau of Economic Research.

Bettinger, Eric P and Bridget Terry Long. 2005. "Do faculty serve as role models? The impact of instructor gender on female students." *American Economic Review* pp. 152–157.

Bird, Karen. 2005. "Gendering parliamentary questions." *The British Journal of Politics & International Relations* 7(3):353–370.

Blei, David and John Lafferty. 2006. "Correlated topic models." *Advances in neural information processing systems* 18:147.

Boothroyd, Baroness Betty. 2013.
URL: *https://www.youtube.com/watch?v=uH6-lfwUeS4*

Brajer, Victor and Andrew Gill. 2010. "Yakity-Yak: Who Talks Back? An Email Experiment*." *Social Science Quarterly* 91(4):1007–1024.

Cameron, A Colin and Douglas L Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources* 50(2):317–372.

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(01):45–68.

Childs, Sarah. 2000. "The New Labour women MPs in the 1997 British Parliament: issues of recruitment and representation." *Women's History Review* 9(1):55–73.

Childs, Sarah. 2004. "A feminised style of politics? Women MPs in the House of Commons." *The British Journal of Politics & International Relations* 6(1):3–19.

Csardi, Gabor and Tamas Nepusz. 2006. "The igraph software package for complex network research." *InterJournal, Complex Systems* 1695(5):1–9.

Dasgupta, Nilanjana and Shaki Asgari. 2004. "Seeing is believing: Exposure to counter-stereotypic women leaders and its effect on the malleability of automatic gender stereotyping." *Journal of Experimental Social Psychology* 40(5):642–658.

Dee, Thomas S. 2007. "Teachers and the gender gaps in student achievement." *Journal of Human Resources* 42(3):528–554.

Eagly, Alice H and Blair T Johnson. 1990. "Gender and leadership style: A meta-analysis." *Psychological bulletin* 108(2):233.

Eggers, Andrew C and Arthur Spirling. 2014. "Ministerial Responsiveness in Westminster Systems: Institutional Choices and House of Commons Debate, 1832–1915." *American Journal of Political Science* 58(4):873–887.

Erkan, Günes and Dragomir R Radev. 2004. "LexRank: Graph-based lexical centrality as salience in text summarization." *Journal of Artificial Intelligence Research* pp. 457–479.

Escobar-Lemmon, Maria and Michelle M Taylor-Robinson. 2008. "Getting to the top: Career paths of women in Latin American cabinets." *Political Research Quarterly* .

Fader, Anthony, Dragomir R Radev, Michael H Crespin, Burt L Monroe, Kevin M Quinn and Michael Colaresi. 2007. MavenRank: Identifying Influential Members of the US Senate Using Lexical Centrality. In *EMNLP-CoNLL*. pp. 658–666.

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 Jennifer L Lawless. 2011. "Gendered perceptions and political candidacies: A central barrier to women's equality in electoral politics." *American Journal of Political Science* 55(1):59–73.

Gilardi, Fabrizio. 2015. "The Temporary Importance of Role Models for Women's Political Representation." *American Journal of Political Science* .

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(3):267–297.

Hannah, Annette and Tamar Murachver. 1999. "Gender and conversational style as predictors of conversational behavior." *Journal of Language and Social Psychology* 18(2):153–174.

Hannah, Annette and Tamar Murachver. 2007. "Gender preferential responses to speech." *Journal of Language and Social Psychology* 26(3):274–290.

Hawkesworth, Mary. 2003. "Congressional enactments of race–gender: Toward a theory of raced–gendered institutions." *American Political Science Review* 97(04):529–550.

Heath, Roseanna Michelle, Leslie A Schwindt-Bayer and Michelle M Taylor-Robinson. 2005. "Women on the sidelines: Women's representation on committees in Latin American legislatures." *American Journal of Political Science* 49(2):420–436.

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(04):583–622.

Imai, Kosuke, Luke Keele, Dustin Tingley and Teppei Yamamoto. 2011. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review* 105(04):765–789.

Karakowsky, Leonard, Kenneth McBey and Diane L Miller. 2004. "Gender, perceived competence, and power displays examining verbal interruptions in a group context." *Small Group Research* 35(4):407–439.

Karpowitz, Christopher F, Tali Mendelberg and Lee Shaker. 2012. "Gender inequality in deliberative participation." *American Political Science Review* 106(03):533–547.

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(03):560–576.

Kleinberg, Jon M. 1999. "Authoritative sources in a hyperlinked environment." *Journal of the ACM (JACM)* 46(5):604–632.

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(03):840–855.

Lockwood, Penelope. 2006. "Someone like me can be successful: Do College Students Need Same-Gender Role Models?" *Psychology of Women Quarterly* 30(1):36–46.

Lovenduski, Joni. 2005. *Feminizing politics*. Polity.

Lovenduski, Joni and Pippa Norris. 1996. *Women in politics*. Vol. 4 Oxford University Press.

Mansbridge, Jane. 1999. "Should blacks represent blacks and women represent women? A contingent "yes"." *The Journal of politics* 61(03):628–657.

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(03):707–733.

Mendelberg, Tali, Christopher F Karpowitz and J Baxter Oliphant. 2014. "Gender inequality in deliberation: Unpacking the black box of interaction." *Perspectives on Politics* 12(01):18–44.

Mendelberg, Tali, Christopher F Karpowitz and Nicholas Goedert. 2014. "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.

Mendez, Jeanette Morehouse and Tracy Osborn. 2010. "Gender and the perception of knowledge in political discussion." *Political Research Quarterly* 63(2):269–279.

Mihalcea, Rada. 2004. Graph-based ranking algorithms for sentence extraction, applied to text summarization. In *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions*. Association for Computational Linguistics p. 20.

Nixon, Lucia A and Michael D Robinson. 1999. "The educational attainment of young women: Role model effects of female high school faculty." *Demography* 36(2):185–194.

O'Brien, Diana Z. 2015. "Rising to the Top: Gender, Political Performance, and Party Leadership in Parliamentary Democracies." *American Journal of Political Science* .

Page, Lawrence, Sergey Brin, Rajeev Motwani and Terry Winograd. 1999. "The PageRank citation ranking: bringing order to the Web.".

Pearson, Kathryn and Logan Dancey. 2011. "Speaking for the underrepresented in the House of Representatives: Voicing women's interests in a partisan era." *Politics & Gender* 7(04):493–519.

Rask, Kevin N and Elizabeth M Bailey. 2002. "Are faculty role models? Evidence from major choice in an undergraduate institution." *The Journal of Economic Education* 33(2):99–124.

Ruddock, Joan. 2003.
URL: *http://www.theyworkforyou.com/debates/?id=2003-03-06.976.0*

Russell, Catherine A and Mark W DeLancey. 2002. "African Women in Cabinet Positions-Too Few, Too Weak: A Research Report." *Asian Women* 15:147–163.

Sapiro, Virginia. 1981. "Research frontier essay: When are interests interesting? The problem of political representation of women." *American Political Science Review* 75(03):701–716.

Studlar, Donley T and Gary F Moncrief. 1999. "Women's work? The distribution and prestige of portfolios in the Canadian provinces." *Governance* 12(4):379–395.

Taylor-Robinson, Michelle M and Roseanna Michelle Heath. 2003. "Do women legislators have different policy priorities than their male colleagues? A critical case test." *Women & Politics* 24(4):77–101.

Thomson, Rob, Tamar Murachver and James Green. 2001. "Where is the gender in gendered language?" *Psychological Science* 12(2):171–175.

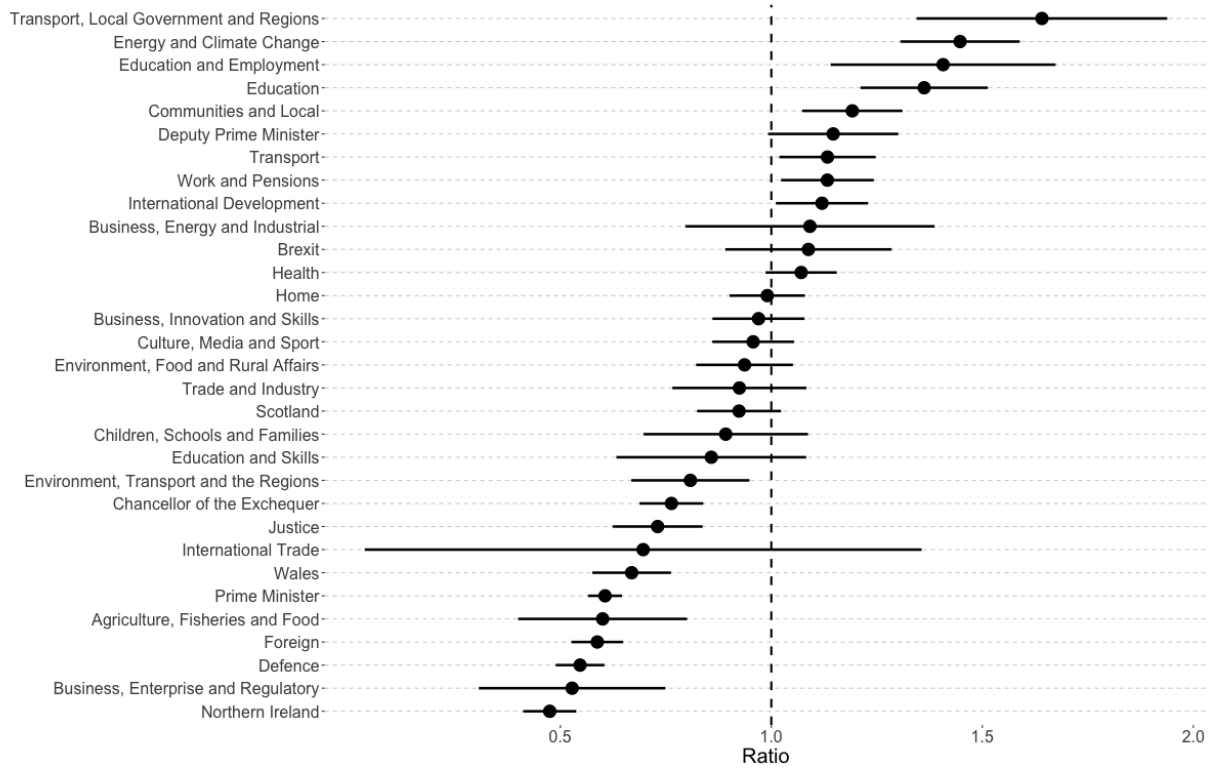
Wang, Mingzhu and Elisabeth Kelan. 2013. "The gender quota and female leadership: Effects of the Norwegian gender quota on board chairs and CEOs." *Journal of business ethics* 117(3):449–466.

Wängnerud, Lena. 2009. "Women in parliaments: Descriptive and substantive representation." *Annual Review of Political Science* 12:51–69.

Wolbrecht, Christina and David E Campbell. 2007. "Leading by example: Female members of parliament as political role models." *American Journal of Political Science* 51(4):921–939.

S1. Appendix

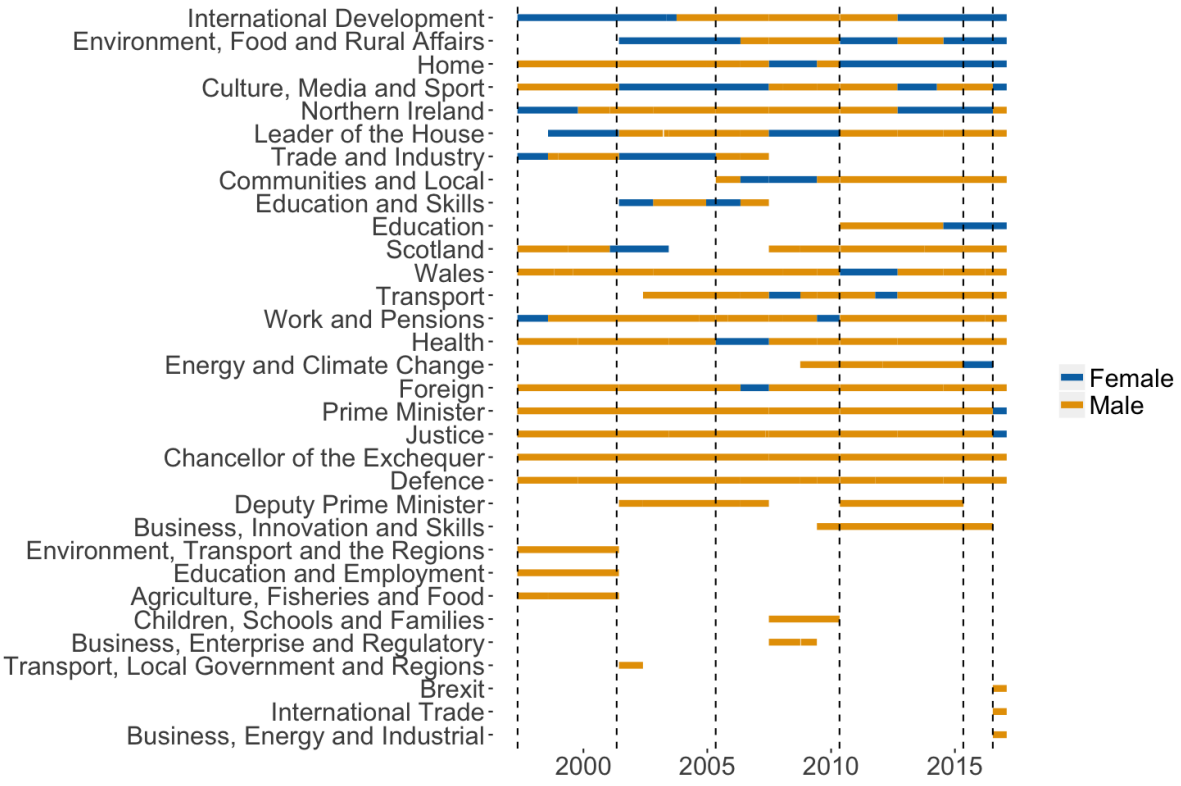
Figure S1: Female speech ratio, by ministry



Note: The figure shows the average female speech ratio as defined in equation S1 for each ministry, pooled across all months in the data. It is clear from the figure that some ministries are subject to greater levels of female participation than others.

A. Independent variable

Figure S2: Gender of ministers over time



Note: The figure shows the distribution of the independent variable over time. While some ministries are never held by a woman (those all in orange), and the Women's ministry is always presided over by a woman (all in blue), the gender of the minister in several ministries varies over time.

Figure S2 shows the variation in the independent variable over time for all 33 ministries included in the sample. Ministries are sorted by the proportion of the time period that the ministry is occupied by a female minister. Orange bars pertain to periods in which the minister responsible is male, and blue bars represent female ministers. There are several ministries for which the responsible minister is never a woman, including the the Defence ministry and the position of Chancellor of the Exchequer.

B. Alternative dependent variables

Equations S1, S2, and S3 provide alternative definitions of the dependent variable in equation 1. Results for the main fixed-effects models using these alternative operationalisations are presented in tables S1, S2, and S3 below. Regardless which of these measures is used,

the main results hold: the appointment of a female minister leads to an increase in the level of participation in parliamentary debates by female MPs.

$$RatioWordsWomen_{dmt} = \frac{PropWordsWomen_{dmt}}{\text{Proportion of women in parliament}_t} \quad \text{(S1)}$$

$$PropSpeechesWomen_{dmt} = \frac{\text{\# speeches by women}_{dmt}}{\text{\# speeches by men and women}_{dmt}} \quad \text{(S2)}$$

$$RatioSpeechesWomen_{dmt} = \frac{PropSpeechesWomen_{dmt}}{\text{Proportion of women in parliament}_t} \quad \text{(S3)}$$

Table S1: Effect of appointing of a female minister on female speech (ratio of words)

	RatioWordsWomen						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female minister	0.290***	0.270***	0.263***	0.227***	0.209***	0.218***	0.169***
	(0.082)	(0.075)	(0.068)	(0.067)	(0.067)	(0.063)	(0.042)
Constant	0.833***	0.539***	0.600***	0.308*	0.652	0.900	0.536
	(0.052)	(0.135)	(0.229)	(0.179)	(8.444)	(2,325.490)	(2.979)
Month FEs	×	✓	×	✓	✓	✓	✓
Ministry FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Effect Size %	35	32	32	27	25	26	20
95% CI	[16,54]	[15,50]	[16,48]	[12,43]	[9,41]	[11,41]	[10,30]
Observations	3,057	3,057	3,057	3,057	3,057	3,057	3,057
R^2	0.034	0.161	0.186	0.299	0.327	0.349	
Adjusted R^2	0.034	0.099	0.178	0.239	0.261	0.278	0.332

NOTE: Models 1-6 represent OLS fixed-effect regressions for the period 1997-2017. Regression coefficients are shown with bootstrapped robust standard errors (clustered by ministry) shown in parentheses. Models 5 and 6 include linear and quadratic ministry-specific time trends in addition to ministry fixed-effects and month fixed-effects. Model 7 is a GAM model including non-parametric, ministry-specific, flexible time trends. The "Effect Size" row indicates the percentage increase in female participation relative to the average female participation rate under male ministers. *p<0.1; **p<0.05; ***p<0.01

Table S2: Effect of the appointment of a female minister on female participation (proportion of speeches)

	Proportion speeches						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female minister	0.043***	0.037***	0.041***	0.030***	0.029***	0.028***	0.022***
	(0.013)	(0.012)	(0.011)	(0.010)	(0.009)	(0.009)	(0.007)
Constant	0.175***	0.075***	0.127**	0.043	0.103	0.459	0.419
	(0.010)	(0.019)	(0.055)	(0.027)	(1.027)	(268.539)	(0.984)
Month FEs	×	✓	×	✓	✓	✓	✓
Ministry FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Effect Size %	25	21	23	17	17	16	13
95% CI	[10,40]	[8,34]	[11,36]	[6,28]	[7,27]	[6,26]	[5,21]
Observations	3,057	3,057	3,057	3,057	3,057	3,057	3,057
R^2	0.023	0.224	0.166	0.336	0.357	0.376	
Adjusted R^2	0.023	0.166	0.158	0.280	0.294	0.307	0.339

Note: Models 1-6 represent OLS fixed-effect regressions for the period 1997-2017. Regression coefficients are shown with bootstrapped robust standard errors (clustered by ministry) shown in parentheses. Models 5 and 6 include linear and quadratic ministry-specific time trends in addition to ministry fixed-effects and month fixed-effects. Model 7 is a GAM model including non-parametric, ministry-specific, flexible time trends. The "Effect Size" row indicates the percentage increase in female participation relative to the average female participation rate under male ministers. *p<0.1; **p<0.05; ***p<0.01

Table S3: Effect of the appointment of a female minister on female participation (ratio of speeches)

	Ratio speeches						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female minister	0.204***	0.186***	0.186***	0.154***	0.153***	0.146***	0.114***
	(0.062)	(0.063)	(0.047)	(0.053)	(0.051)	(0.049)	(0.038)
Constant	0.865***	0.412***	0.696***	0.237*	0.570	2.234	2.342
	(0.046)	(0.110)	(0.178)	(0.139)	(4.512)	(1,181.458)	(5.565)
Month FEs	×	✓	×	✓	✓	✓	✓
Ministry FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Effect Size %	24	22	22	18	18	17	13
95% CI	[9,38]	[7,36]	[11,32]	[6,30]	[6,29]	[6,28]	[5,22]
Observations	3,057	3,057	3,057	3,057	3,057	3,057	3,057
R^2	0.022	0.136	0.155	0.258	0.279	0.299	
Adjusted R^2	0.022	0.072	0.146	0.194	0.208	0.222	0.258

Note: Models 1-6 represent OLS fixed-effect regressions for the period 1997-2017. Regression coefficients are shown with bootstrapped robust standard errors (clustered by ministry) shown in parentheses. Models 5 and 6 include linear and quadratic ministry-specific time trends in addition to ministry fixed-effects and month fixed-effects. Model 7 is a GAM model including non-parametric, ministry-specific, flexible time trends. The "Effect Size" row indicates the percentage increase in female participation relative to the average female participation rate under male ministers. *p<0.1; **p<0.05; ***p<0.01

C. Details of the influence measurement procedure

I proceed in two steps: first, I construct similarity graphs for all speeches in each debate; second, I analyse the graphs using an iterative ranking algorithm to calculate a vector of centrality scores, P, which correspond to the influence of each speech in each debate.

Construction of a debate-specific similarity graph, S_d, begins with the selection of a metric which measures how linguistically similar two speeches are to one another. I represent each speech as an N-dimensional term-frequency-inverse-document-frequency (tf-idf) vector, where N is the number of unique words in the corpus. Each element in the vector is a count of the number of times a given word, w, appears in a given speech, s, multiplied by the logged *inverse document frequency* of that word, to create a weighted term-frequency score, v_{ws}, for each word in each speech. Where N is the total number of unique words in the corpus, n_w is the number times that word w appears in the corpus, and tf_{ws} is the number times that word w appears in speech s, the score for w in s is given by:

$$v_{ws} = tf_{ws} * log(\frac{N}{n_w}) \tag{S4}$$

A high value of v_{ws} occurs when a word is used frequently in a given speech, but infrequently in the corpus as a whole. The weights thus filter out very common words such as 'stopwords', and ensure that the vector representation of the speeches mostly reflects topically-salient features of the political debate.

Having calculated the tf-idf vectors for each speech in the corpus, I construct D similarity matrices (one for each debate), the typical element of which is: $S_d(i,j) = sim(v_i, v_j)$, the cosine-similarity of the weighted word-count vectors of speeches i and j in debate d. Each graph (again, one for each debate) therefore consists of nodes that represent speeches in a debate, and edges which are placed between speeches for which $sim(v_i, v_j)$ is greater than some threshold value, S_{min}.[11] The edges are then weighted by the similarity scores.

The cosine-similarity relation is symmetric (i.e. because $S_d(i,j) = S_d(j,i)$) and thus it is possible to construct either undirected (where edges between nodes run in both directions and receive the same weight) or directed (where edges between nodes run in only one direction) networks (Erkan and Radev, 2004). As I conceptualise influence as the degree to which language used in one speech is adopted in subsequent speeches, it is necessary to take the temporal ordering of debate into account when constructing the graphical network. Put simply, it does not make sense for speeches that occur later in the debate to 'influence' speeches that occur earlier in the debate. I therefore focus on only the upper triangle of the similarity matrices, S_d, while setting all elements in the lower triangle to zero. The consequence of this is that 'references' from one speech to another can only flow in one direction: later speeches can reference earlier ones, but not vice versa. Using a directed graph makes no difference to the computation of the influence scores (Mihalcea, 2004).

As described above, the influence of a speech is determined by the number of references it receives from other speeches within a debate (i.e. by the number of speeches which

[11] Throughout the analysis I set S_{min} to 0.25, in line with Fader et al. (2007).

are linguistically similar to it), and by the influence of the referencing speeches. In the simplified case where all edges receive a weight of 1, an intuitive way of formulating this idea is to imagine that each speech has an influence value, and that this value to gets distributed to the speeches that it references:

$$p(i) = \sum_{j \in adj(i)} \frac{p(j)}{deg(j)} \tag{S5}$$

Where $p(i)$ is the influence of speech i, $adj(i)$ is the set of speeches that have edges with i, and $deg(j)$ is the degree of node j (the degree of a node is simply the number of edges that connects the node to other nodes). This formulation emphasises that a speech is more influential when it is referenced by many other speeches ($adj(i)$), when the influence of the referencing speeches ($p(j)$) increases, and when the referencing speeches reference relatively few other speeches ($deg(j)$). Weighting the edges of the network by $S_d(i, j)$ allows references to vary in strength (according to the similarity between speeches i and j) and we can reformulate equation S5 to include the weights in S_d via:

$$p(i) = \sum_{j \in adj(i)} \frac{S_d(i, j)}{\sum_{k \in adj(j)} S_d(k, j)} p(j) \tag{S6}$$

Equation S6 makes clear that the reference that speech i receives from speech j is determined by the linguistic similarity between i and j (the numerator), and the similarity between j and all of the speeches that j references (the denominator). Fader et al. (2007) and Erkan and Radev (2004) show that computation of the vector of speech-level influence scores, P, is achieved by calculating the left eigenvector of the row-normalised similarity matrix S_d via the *PageRank* algorithm, which was originally designed for computing webpage prestige in the Google search engine (Page et al., 1999).[12] Finally, with these speech-influence scores in hand, as some MPs will speak multiple times in a debate, the influence score of an MP in a given debate is simply the sum of the influence scores p for the speeches given by that MP in that debate.

D. HITS algorithm influence results

This section replicates the analysis in table 2 and figure 4 estimating P in equation S6 using Kleinberg's (1999) HITS algorithm. As table S4 and figure S3 show, the results are substantively and statistically very similar regardless of the estimation approach.

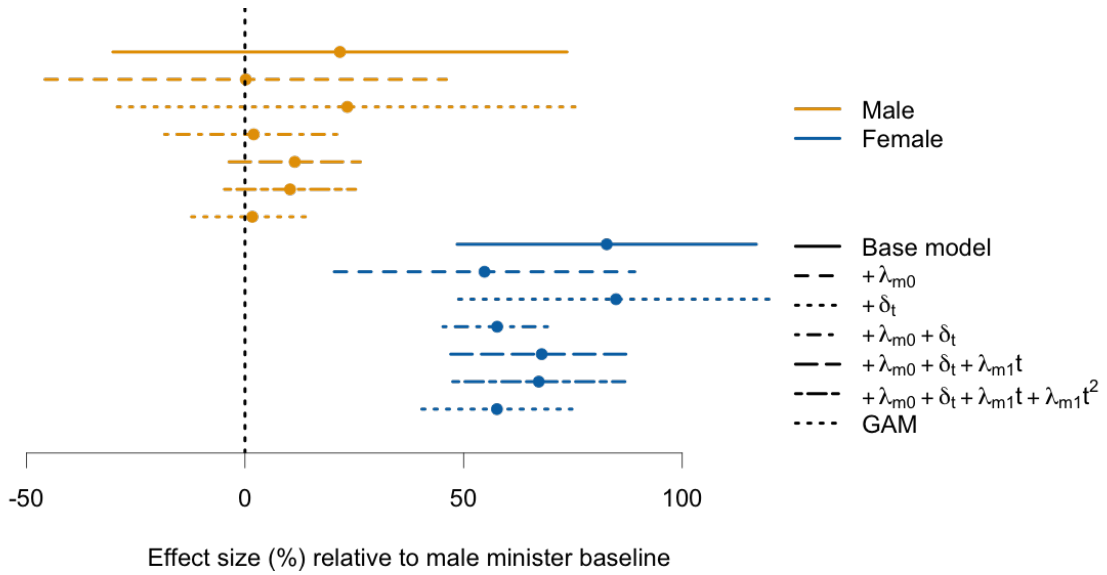
[12] Mihalcea (2004) shows that either the Kleinberg (1999) HITS algorithm and PageRank can be used to calculate P and that both perform well in approximating human judgements. I present statistically and substantively similar results from the HITS algorithm in appendix section D. I implement both algorithms using the iGraph package in *R* (Csardi and Nepusz, 2006)

Table S4: Effect of appointing a female minister on MPs' debate influence – HITS

	Influence						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	0.222***	0.257***	0.400***	0.333***	0.363***	0.372***	−8.797
	(0.024)	(0.001)	(0.036)	(0.021)	(0.034)	(0.027)	(16.571)
Female minister	0.023	0.0002	0.025	0.002	0.012	0.011	0.002
	(0.028)	(0.025)	(0.028)	(0.011)	(0.008)	(0.008)	(0.008)
Female MP	−0.016**	−0.015*	−0.003	−0.007	−0.007	−0.007	−0.007*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.004)
Interaction	0.070***	0.061***	0.070***	0.063***	0.064***	0.064***	0.063***
	(0.018)	(0.012)	(0.018)	(0.013)	(0.013)	(0.013)	(0.008)
Ministry FEs	×	✓	×	✓	✓	✓	✓
Month FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Observations	174,419	174,419	174,419	174,419	174,419	174,419	174,419
R^2	0.001	0.024	0.013	0.033	0.034	0.035	
Adjusted R^2	0.001	0.024	0.012	0.032	0.033	0.033	0.039

NOTE: Models 1-6 present OLS fixed-effect regressions for the period 1997-2017. Regression coefficients are shown with robust standard errors (clustered by ministry) shown in parentheses. *p<0.1; **p<0.05; ***p<0.01

Figure S3: Marginal effect of female minister on male and female influence – HITS



NOTE: The plot shows the marginal effect of the appointment of a female cabinet minister on the debate influence of male (orange lines) and female (blue lines) MPs, relative to the average level of influence when the minister is male.

E. VALIDATING INFLUENCE SCORES

I test two relatively unambiguous intuitions about which actors in the House of Commons we expect to be influential in parliamentary debate.[13] First, government ministers should be on average *more* influential than other MPs when participating in plenary debate. Ministers play a crucial role in setting the agenda for parliamentary business, and their speeches are frequently used to outline policy that we would expect others to comment on extensively. Second, the Speaker of the House should be on average *less* influential than other members. The majority of the Speaker's contributions are procedural, having little to do with the substantive matters under discussion, and should not be referenced frequently by other members. I test these expectations by regressing the influence score on binary indicators for whether an MP is either the cabinet minister responsible for the current debate, or the Speaker. The results, presented in table S5 in the appendix, strongly support the expectations: cabinet ministers are on average 5 times more influential than backbench MPs, while the Speaker is less than half as influential as backbenchers.

In addition, we can also compare the influence score of an MP to how many direct references that MP receives during a debate. MPs follow strict conventions when directly referring to other members as they must not use the names of their colleagues, but instead refer to the "Member for Holborn and St Pancras" or the "Honourable Member for Taunton

[13]Fader et al. (2007) have shown previously that these influence scores calculated for speeches made in the US Senate correlate strongly with membership and seniority in Senate legislative committees.

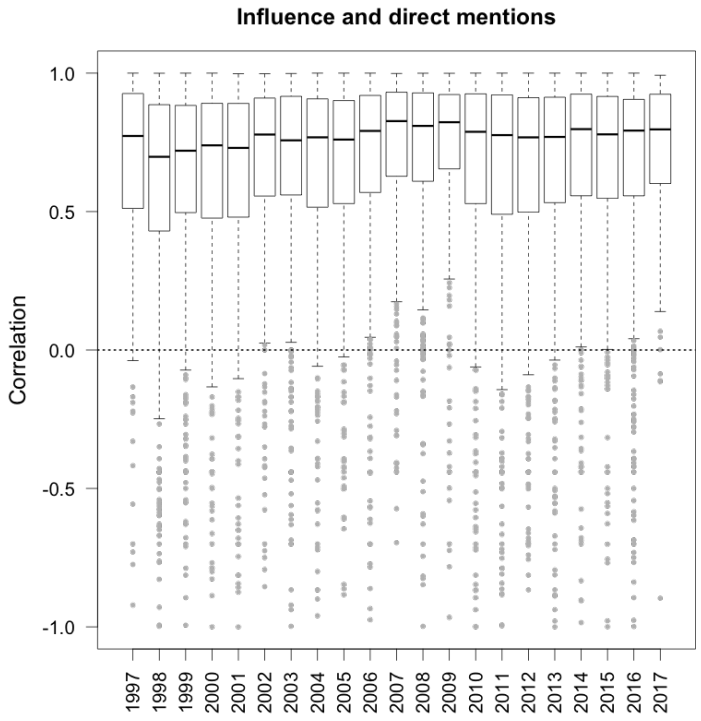
Table S5: Ministers are more influential, and Speakers of the House are less influential

	Influence		
	(1)	(2)	(3)
Minister	0.958***		0.952***
	(0.006)		(0.006)
Speaker		−0.219***	−0.147***
		(0.008)	(0.008)
Constant	0.221***	0.299***	0.226***
	(0.002)	(0.002)	(0.002)
Observations	195,828	195,828	195,828
R^2	0.127	0.003	0.129

NOTE: OLS regressions where the outcome variable is *influence* as defined in equation S6, and the independent variables are "Minister" – an indicator that is equal to one when a speech is given by a government minister – and "Speaker" – an indicator that is equal to one when a speech is given by the Speaker of the House. The baseline corresponds to the average level of *influence* for speeches delivered by backbench MPs. *p<0.1; **p<0.05; ***p<0.01

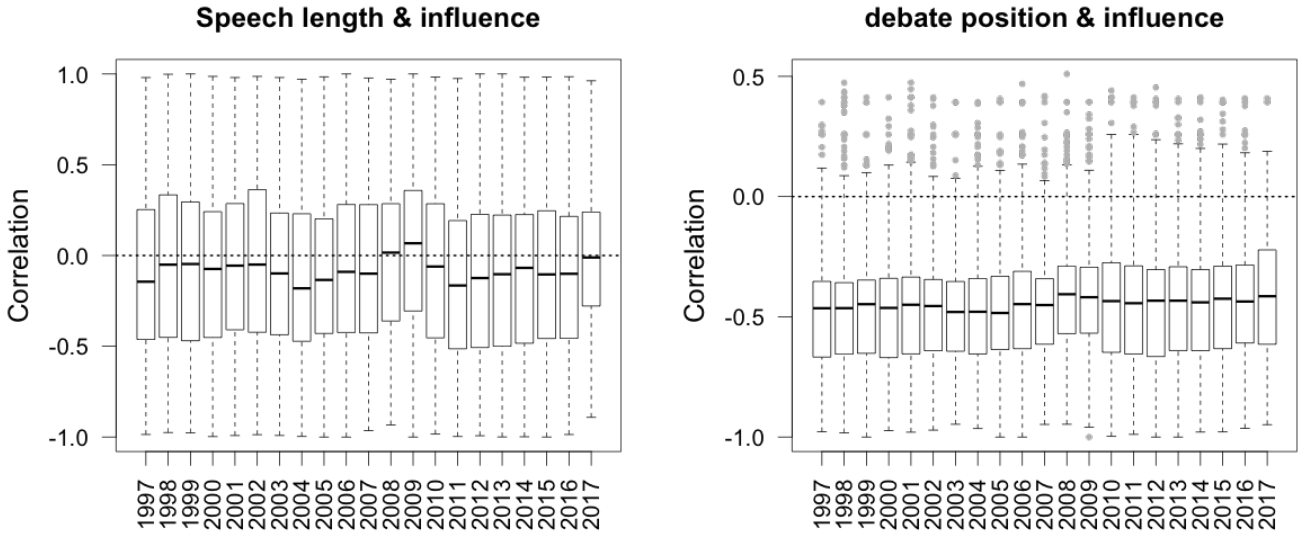
Deane" and so on. Constituency names are unique to each MP, and by searching for constituencies in the speech-texts it is possible to construct a count for the number of times any particular MP is directly mentioned by any other MP during the course of a debate. It seems clear that an MP who is directly mentioned by many other MPs in their speeches is playing an important role in the debate at hand and if the influence score defined in equation S6 is valid, it should correlate positively with the number of direct mentions that an MP receives in debate. In appendix figure S4 I show that this is the case: the average correlation across all debates in the sample is 0.66. Together, these comparisons provide reassuring evidence regarding the face validity of the measure of influence described above.

Figure S4: Correlation of speech 'influence' and direct mentions



NOTE: The plot shows the average correlation between the influence of an MP and the number of times that that MP was directly mentioned by other members measured at the debate level across all years in the sample. In line with expectations, influence is strongly positively correlated with direct mentions.

Figure S5: Correlation of speech 'influence' with speech length and debate position



NOTE: The left panel shows the average correlation between the length of a speech and influence measured at the debate level across all years in the sample. The right panel shows the equivalent correlation between speech position and influence.

F. Validating responsiveness scores

First, within a debate, MPs might use similar words even when they are not responding to one another. Debates are normally focussed on a small number of topics, the discussion of which will lead MPs to use similar language regardless of whether they are talking directly to one another. However, if the measure defined in equation 6 captures responsiveness, and not merely topicality, then speeches that are adjacent to one another should demonstrate higher responsiveness scores than speeches that are not adjacent. Table S6 in the appendix tests this hypothesis. From each debate in the corpus, I randomly sample two speech pairs. One of the pairs is adjacent, and one is non-adjacent. I then regress the responsiveness score on a binary indicator which is equal to one for an adjacent pair of speeches, and zero otherwise. The coefficient on this indicator is statistically significant, and implies that adjacent speeches are approximately 30% more responsive than non-adjacent speeches. This provides strong evidence that equation 6 is capturing something distinct from topicality: comparing pairs of speeches *within the same debate*, those speeches that follow directly after each other are more responsive than speeches that are non-adjacent.

Second, in a subset of debates, government ministers go before the House to field questions from backbenchers, and are required to provide answers to these questions. In these 'Question Time' debates, questions by backbenchers need not address the same topic as the question just answered by the minister. For example, a first backbencher might ask the

Table S6: Adjacent speeches are more responsive than non-adjacent speeches

	res
Adjacent	8.133***
	(1.163)
Constant	27.898***
	(0.823)
Observations	21,604
R^2	0.002

NOTE: OLS regression for adjacent and non-adjacent speeches (within a debate). Regression coefficients are shown with standard errors in parentheses. The outcome variable is *res* as defined in equation 6, and the independent variable is an indicator that is equal to one when a given pair of speeches occupy adjacent positions in the debate. The baseline corresponds to speeches that are non-adjacent. *p<0.1; **p<0.05; ***p<0.01

minister about schools, to which the minister will provide an answer, and then a second backbencher might ask about child care provision, to which the minister must also respond. In these debates, we should therefore expect that when a minister's speech follows a backbencher, that speech should be more responsive than when a backbencher's speech follows that of a minister. For each speech in each 'Question Time' debate, I code whether the speech is made by a minister responding to a backbencher, or a backbencher asking a new question.[14] I then regress the responsiveness score on a binary indicator which is equal to one when the speech is made by a minister in response to a backbencher. The results in table S7 show that ministerial replies are more than twice as responsive than are questions posed by backbenchers to the minister. This indicates that the measure is accurately recovering intuitive properties of the concept of responsiveness.

[14]I exclude all instances where a backbencher follows from another backbencher.

Table S7: Minister and backbencher responsiveness

	res
Minister responding to backbencher	10.036***
	(0.167)
Constant	10.028***
	(0.121)
Observations	152,683
R^2	0.023

NOTE: OLS regressions for 9927 "question time" debates. Regression coefficients are shown with standard errors in parentheses. The outcome variable is *res* as defined in equation 6. The independent variable is an indicator that is equal to one when a speech is spoken by a minister, and comes immediately after a speech by a backbencher. The baseline corresponds to backbench speeches that follow directly after a speech by a minister. *p<0.1; **p<0.05; ***p<0.01

G. Strategic appointment of female shadow ministers

To investigate the hypothesis that opposition parties may respond strategically to the appointment of a female government minister by appointing a female shadow minister to the opposition cabinet, I analyse the relationship between the sex of a *newly appointed shadow minister* and the sex of the *current government minister*. I estimate this relationship using logit models of the following form:

$$\text{logit}(\mathbb{E}[ShadowFemaleMinister_{mt}]) = \alpha + \beta_1 * FemaleMinister_{mt} + \lambda_m + \epsilon_{mt}$$

where $ShadowFemaleMinister_{mt}$ is equal to one when the shadow minister appointed to ministry m at time t is a woman, and zero otherwise. $FemaleGovernmentMinister$ is equal to one when the government minister responsible for a given ministry m in time t is a woman, and zero otherwise. λ_m is a ministry fixed effect. If opposition parties are responding strategically to the sex of the government minister in a given ministry, then the β_1 coefficient will be positive, indicating that the probability of appointing a female shadow minister is associated with the sex of the current cabinet minister for that ministry. The results of these regressions are given in table S8.

While the coefficient on the 'female government minister' variable are positive in both models, these effects are imprecisely estimated, and statistically indistinguishable from zero. This suggests that it is unlikely that the effects documented in the main analysis are driven by the strategic appointment of female shadow ministers by opposition parties.

Table S8: Probability that a female shadow minister will be appointed, conditional on the sex of the current government minister

	Female shadow minister	
	(1)	(2)
Female government minister	0.141	0.504
	(0.456)	(0.659)
Constant	-1.288^{***}	-0.504
	(0.399)	(1.560)
Ministry FEs	No	Yes
Observations	149	149
Note:	*p<0.1; **p<0.05; ***p<0.01	
	Logit model	

H. Differential agenda-setting of male and female ministers

The main idea here is to measure the topical content of the issues under discussion in debate, and to evaluate whether topics which are associated with high levels of female participation (when the minister is male) increase when a female minister is appointed. In order to measure the topical content of the legislation under debate, I focus on the speeches made by *ministers* during each debate, rather than the speeches made by all members. In many cases, debates begin with a long opening statement by the minister, in which they put forward the purpose and detail of the legislation to be considered by the House. As the content of speeches made by other members may itself be a result of the appointment of a female minister, these speeches provide a useful resource for estimating the agenda proposed by the ministers.

I proceed in four steps. First, I estimate a series of topic models to produce debate-level topic proportions for all debates in the sample. These proportions indicate the topical content of each debate, and give a basis on which to find thematically similar debates under both male and female ministers. Second, I use the topic proportions for debates which are held under male ministers as explanatory variables in linear regressions, where the dependent variable is the ratio of words spoken by women as defined in equation S1. The coefficients from these regressions indicate the degree to which each latent topic is traditionally associated with female participation in debate. Third, in a second set of linear regressions, I estimate the relationship between the prevalence of a topic and the sex of a minister. The coefficients from these regressions indicate whether a topic increases or decreases when the minister is female. Finally, I compare the two sets of regression coefficients. If the agenda-setting hypothesis is correct, there should be a positive correlation between these two sets of coefficients: topics that are traditionally associated

Table S9: Effect of appointing a female minister on female debate participation – shadow ministers excluded

	PropWordsWomen						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female minister	0.035**	0.034**	0.037***	0.035***	0.043***	0.036***	0.031***
	(0.015)	(0.014)	(0.014)	(0.013)	(0.008)	(0.011)	(0.007)
Constant	0.169***	0.107***	0.135**	0.085**	0.107	−0.043	0.076
	(0.006)	(0.025)	(0.052)	(0.038)	(1.162)	(160.703)	(0.181)
Month FEs	×	✓	×	✓	✓	✓	✓
Ministry FEs	×	×	✓	✓	✓	✓	✓
Linear time trends	×	×	×	×	✓	✓	×
Quadratic time trends	×	×	×	×	×	✓	×
Flexible time trends	×	×	×	×	×	×	✓
Effect Size %	21	20	22	21	25	21	18
95% CI	[4,38]	[4,37]	[6,38]	[5,36]	[16,35]	[8,34]	[10,27]
Observations	3,234	3,234	3,234	3,234	3,234	3,234	3,234
R^2	0.014	0.119	0.135	0.231	0.256	0.281	
Adjusted R^2	0.013	0.057	0.126	0.169	0.187	0.206	0.224

NOTE: Models 1-6 represent OLS fixed-effect regressions for the period 1997-2017 with shadow ministers excluded from the data. Regression coefficients are shown with bootstrapped cluster-robust standard errors (clustered by ministry) shown in parentheses. *p<0.1; **p<0.05; ***p<0.01

with female participation will increase when the minister is female. Such a finding would suggest that female ministers are indeed focussing on topics that are more conducive to female participation in legislative debate.

I start by applying a series of unsupervised topic models to all speeches made by ministers in the entire sample. I use the Correlated Topic Model (Blei and Lafferty, 2006), which, as with all topic models, assumes that the frequency with which terms co-occur within different documents (here, debates) gives information about the topics that feature in those documents. The key quantity of interest recovered from the CTM is θ, which is a T x D matrix of topic proportions that describe the fraction of each ministerial statement $d \in \{1, 2, ..., D\}$ that is from each topic $t \in \{1, 2, ..., T\}$. Analysts must choose how many topics to estimate from the data, and because the 'correct' number of topics is unclear, *a priori*, I estimate K topic models for a range of topic counts from 20 to 60, at 10 topic increments. This results in 5 separate θ_k matrices, with typical elements θ_{ktd}: the proportion of ministerial-statement d in topic t from topic-model k.

I then use each θ_k matrix as the model matrix[15] for a linear regression predicting Y_d, the female speech ratio in debate d. As the goal of this first-stage model is to establish a baseline level of female participation associated with each topic, I estimate this model only for those debates where the presiding minister is male. I repeat this exercise K times, once for each topic model.

$$Y_d = b_{k1}\theta_{k1d} + b_{k2}\theta_{k2d} + ... + b_{kT}\theta_{kTd} + \epsilon_d \tag{S7}$$

The estimated b coefficients represent the degree to which each topic (collection of words) is associated with female participation in debates, holding other topics constant. An example of the substantive information that these coefficients contain is clear from table S10 , which contains each topic from the 30 topic model, ordered by their respective b coefficients. Reassuringly, the topics with the largest b coefficients deal primarily with topics that match intuitive notions of female interests, including children, parents, and women's issues. Additionally, women appear relatively more likely to contribute to debates that focus on the NHS, teachers and schools, and energy issues.

Next, I estimate a series of regressions to establish which topics are more prevalent under female ministers. As we are concerned here with establishing the differences in agenda-setting *within* government ministries, I estimate models of the following form:

$$\theta_{ktd(m)} = \alpha + \gamma_{kt} * FemaleMinister_d + \lambda_m + \epsilon_{ktd} \tag{S8}$$

Where $\theta_{ktd(m)}$ is the proportion of debate-text d (in ministry m) devoted to topic t from topic model k. $FemaleMinister_d$ is a binary variable equal to one when debate d is presided over by a female minister, and λ_m is a ministry fixed effect. The model is estimated separately for each *topic*, and, as in the previous step, I repeat this exercise for each of the 5 *topic models*. The estimation therefore results in K vectors of γ_t coefficients – one coefficient for

[15]The topic proportions for each statement (θ_{kd}) sum to one, and so I could exclude one of the topics or the intercept term. I choose to exclude the intercept term.

Table S10: Topics ordered by their association with female participation under male ministers

Topic	beta
child_women_children_disabl_parent_famili_employ	1.330
nhs_patient_clinic_doctor_hospit_gps_trust	0.968
speech_opposit_chamber_bench_conserv_liber_lord	0.944
debat_motion_monday_thursday_committe_tuesday_wednesday	0.834
local_author_flood_council_grant_area_fund	0.789
look_sure_absolut_import_think_see_right	0.766
carbon_energi_emiss_gas_climat_electr_renew	0.758
rail_railway_passeng_transport_railtrack_airport_road	0.677
teacher_school_pupil_student_educ_teach_academi	0.653
bank_deficit_labour_tax_budget_chancellor_economi	0.582
offend_sentenc_prison_crime_polic_offenc_victim	0.578
role_bbc_process_approach_white_relationship_detail	0.569
refuge_humanitarian_zimbabw_syria_pakistan_syrian_isil	0.545
sport_olymp_scottish_scotland_footbal_game_lotteri	0.522
amend_claus_bill_committe_provis_legisl_draft	0.451
post_compani_manufactur_busi_industri_enterpris_small	0.443
pension_incom_credit_minimum_retir_payment_earn	0.441
health_servic_treatment_wait_cancer_nurs_extra	0.375
inquiri_investig_report_sir_alleg_inform_recommend	0.366
farmer_anim_farm_vaccin_outbreak_diseas_agricultur	0.363
immigr_asylum_court_appeal_terror_card_ident	0.360
global_aid_develop_world_india_africa_debt	0.355
saddam_palestinian_iran_israel_iraq_resolut_weapon	0.200
troop_afghanistan_afghan_armi_defenc_deploy_nato	0.167
wale_welsh_assembl_elector_devolut_elect_referendum	0.108
think_tri_thing_say_reason_let_problem	0.065
treati_europ_european_union_negoti_enlarg_britain	0.026
ireland_northern_paramilitari_irish_agreement_decommiss_ira	-0.004
cent_per_conserv_cut_spend_billion_rise	-0.166
awar_matter_certain_understand_although_concern_gentleman	-0.272

each topic, in each topic model. When γ_{kt} is positive, this implies that the use of the topic increases when a female minister is appointed, and when it is negative it suggests that the use of the topic decreases on the appointment of a female minister.

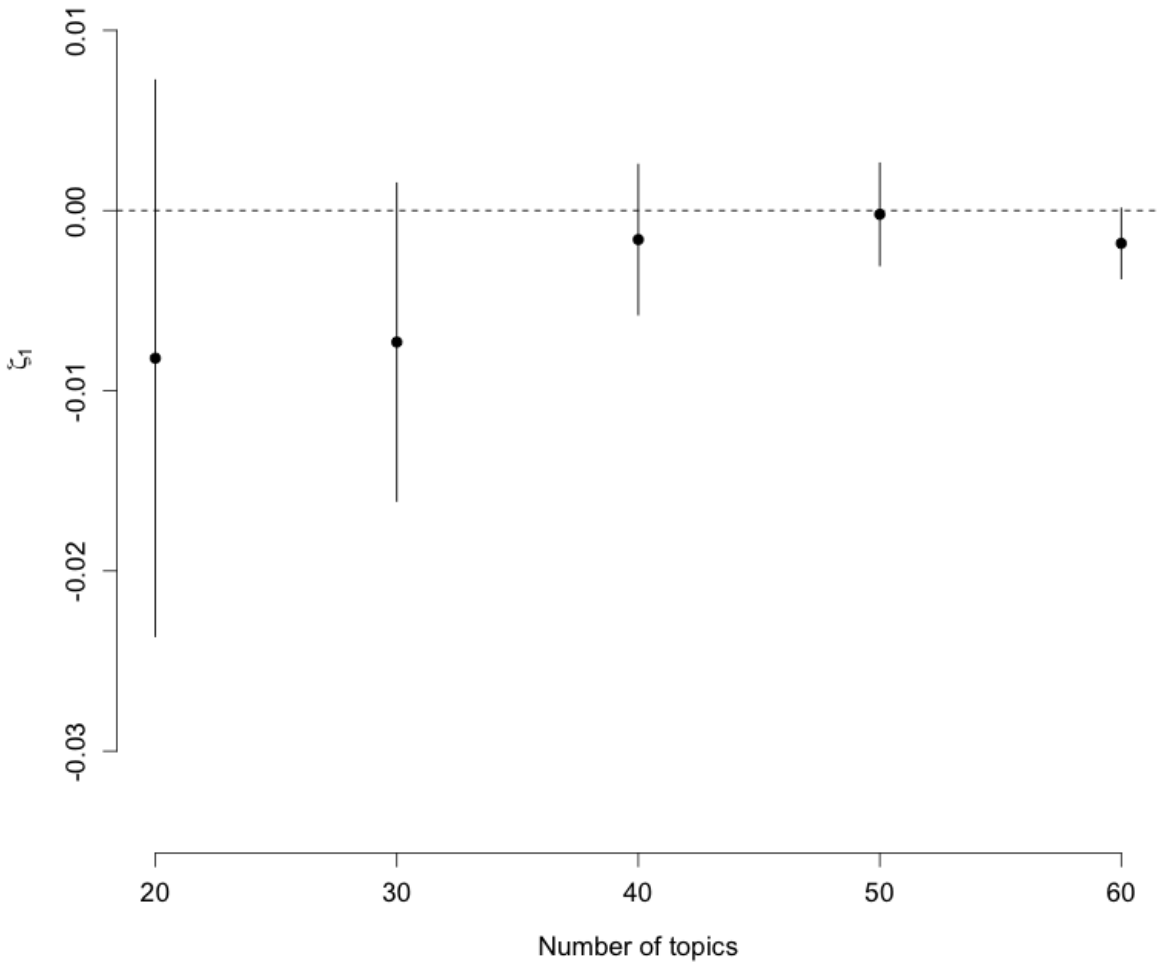
Equations S7 and S8 therefore result in two vectors of coefficients: b_k gives the relationship between each of the topics in topic model k and the level of female debate participation under male ministers, and γ_k indicates how much each of the same topics in topic model k increases (or decreases) when a female minister is appointed. Assessing the correlation between these coefficient vectors allows us to test whether female ministers introduce legislation that focusses on topics which are associated with high levels of female participation under male ministers. Thus, to test the agenda-setting hypothesis, I regress the estimated b coefficients from equation S7 on the γ coefficients from equation S8 according to:

$$b_{t(k)} = \alpha + \zeta_k * \gamma_{t(k)} + \epsilon_{t(k)} \tag{S9}$$

If the agenda-setting hypothesis is correct, then the $b_{t(k)}$ and $\gamma_{t(k)}$ coefficients should be positively correlated, indicating that high female-participation topics (under male ministers) play a more prominent role on the policy agenda when a female minister is appointed. That is, we expect the ζ_k coefficient from equation S9 to be positive. Such a finding would contradict the legislative role-model hypothesis, as it would suggest that the increased levels of legislative participation documented in the main results section could be attributed to the development of an increasingly 'female-friendly' agenda under female ministers. I present the estimated ζ_k coefficients – one for each of the topic models – in figure S6 along with their associated 95% confidence intervals.

Figure S6 provides no clear evidence that female ministers focus more attention on topics that are traditionally marked by high levels of female participation. As none of the slopes is statistically significant at traditional levels, the plot suggests that it is unlikely that changes to the legislative agenda are responsible for the changes in female participation documented in the main text.

Figure S6: There is no increase in the use of 'female friendly' topics on the parliamentary agenda when a female minister is appointed.



NOTE: The graph plots, on the x-axis, the number of topics, and on the y-axis, the estimated ζ coefficients from equation S9. There is no clear evidence that when female ministers are appointed, they focus more on topics that are traditionally popular with other female MPs.