Appendix

Table of contents

A. Model - Proof and Additional Analyses

1. Proof of Equation 3

Lemma 1. *The reinforcement learning model described in Section 3 defines a stochastic process for* $\left(V_{i,m}^{GI}, V_{i,m}^{other}\right)_{m\geq 1}$ *that has a unique stationary distribution characterized by the following density:*

$$h\left(V^{GI}, V^{other}\right) = e^{-\frac{r^2\sigma^2\gamma}{2(2-\gamma)}} \frac{e^{-rV^{GI}-\pi_i^{GI}} + e^{-rV^{other}}}{e^{-r\mu_i^{GI}-\pi_i^{GI}} + e^{-r\mu_i^{other}}} g_i^{GI}(V^{GI}) g_i^{other}(V^{other})$$

where, for $k \in \{GI, other\}$, $g_i^k(\cdot)$ *is a normal density with mean* $\mu^k{}_i$ *and variance* $\sigma^2\gamma/(2-\gamma)$.

Proof. This follows from Lemma 2 in Le Mens et al. (2019).

Proof. The asymptotic probability of choosing the gender topic is obtained by integration of the choice probability (equation 1) with respect to the joint asymptotic density described in Lemma 1. (12)

$$\begin{aligned} A_\infty^{GI} &= \int_{V^{GI}, V^{other}} \frac{1}{1 + e^{-(\pi_i^{GI} + r(V^{GI} - V^{other}))}} dV^{GI} dV^{other} \\ &= \frac{e^{-\frac{r^2\sigma^2\gamma}{2(2-\gamma)}}}{e^{-r\mu_i^{GI}-\pi_i^{GI}} + e^{-r\mu_i^{other}}} \int_{V^{GI}, V^{other}} e^{-rV^{other}} g_i^{GI}(V^{GI}) g_i^{other}(V^{other}) dV^{GI} dV^{other} \\ &= \frac{e^{-\frac{r^2\sigma^2\gamma}{2(2-\gamma)}}}{e^{-r\mu_i^{GI}-\pi_i^{GI}} + e^{-r\mu_i^{other}}} \int_{V^{other}} e^{-rV^{other}} g_i^{other}(V^{other}) dV^{other}. \end{aligned}$$

Noting that $\int_{V^{other}} e^{-rV^{other}} g_i^{other}(V^{other}) dV^{other}$ is the moment generating function of the distribution $g_i^{other}(\cdot)$, evaluated at $-r$, we have:

$$\int_{V^{other}} e^{-rV^{other}} g_i^{other}(V^{other}) dV^{other} = e^{-r\mu_i^{other} + \frac{r^2\sigma^2\gamma}{2(2-\gamma)}}.$$

We finally obtain:

$$\begin{aligned} A_\infty^{GI} &= \frac{e^{-r\mu_i^{other}}}{e^{-r\mu_i^{GI}-\pi_i^{GI}} + e^{-r\mu_i^{other}}} \\ &= \frac{1}{1 + e^{-(\pi^{GI} + r(\mu_i^{GI} - \mu_i^{other}))}}. \end{aligned}$$

□

2. Alternative Specifications of the Valuation Updating Rule

The model analyzed in the body of the paper assumes that the feedback weight in the valuation updating equation (γ in eq. 2) is constant. This is as if earlier feedback instances receive a lower weight than the more recent feedback instances. Here, we relax this assumption and analyze what happens when all past feedback instances receive the same weight. In other words, we assume that the valuation of issue k is the arithmetic average of all feedback instances obtained about this issue (we assume the initial valuation to be a random draw from the feedback distribution – as such initial valuations generally differ for the two issues). In this case the law of large number implies that, in the long run, the valuation of issue k, $V_{i,m}^{k}$ almost surely converges to the mean of the payoff distribution μ^{k}_{i}. Therefore, issue attention almost surely converges to the same quantity as with the model analyzed in the paper (see equation 3).

It is also possible to analyze a 'rational' model in which the agent possesses correctly specified priors and valuations are updated according to Bayes' rule. More specifically, let f^{GI} denote the prior on the mean μ^{GI}_{i}. Similarly, let f^{other} denote the prior on the mean μ^{other}_{i}. We assume that means are realizations of the priors: $\mu^{GI}_{i} \sim f^{GI}$, $\mu^{other}_{i} \sim f^{other}$.

Let $B_{i,m}^{GI}$ denote the mean of the posterior distribution on the payoff of GI at the time of posting message m. We define the probability that agent i chooses issue k for message m in terms of $B_{i,m}^{GI}$

and $B_{i,m}^{other}$:

$$A_{GIi,m} = Logit(\pi_{iGI} + r\Delta B_{i,m}),$$

where $\Delta B_{i,m} = B_{i,m}^{GI} - B_{i,m}^{other}$.

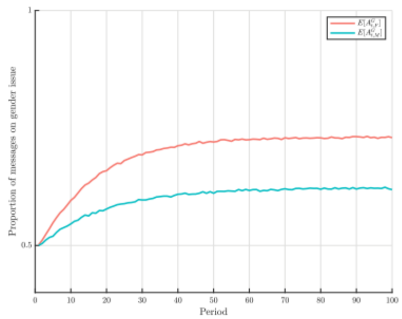
$(B_{i,m}^{GI})_{m\geq 1}$ is a discrete-time martingale, and $E[B_{i,m}^{GI}] = E[\mu^{GI}_{i}]$ for all $m \geq 1$. Similarly, $E[B_{i,m}^{other}] = E[\mu^{other}_{i}]$ for all $m \geq 1$. Moreover, the optional sampling theorem implies these equalities hold asymptotically: $E[B_{i,\infty}^{GI}] = E[\mu_i^{GI}]$ and $E[B_{i,\infty}^{other}] = E[\mu_i^{other}]$.

For simplicity, let us assume that the priors on the means are normally distributed and that the payoffs distributions are normal (as in the paper). In this case, the asymptotic posterior for agent *i* and issue *k* is a single value distribution equal to μ^k_i (the variance of the posterior converges to 0). The asymptotic attention to gender issues for agent *i* is thus the same as with model analyzed in the body of the paper (eq. 3). This implies that our main results about issue attention are not contingent on assuming that agents engage in biased processing of information or have mistaken priors. They are also produced by rational information processing of possibly un-representative samples of information.

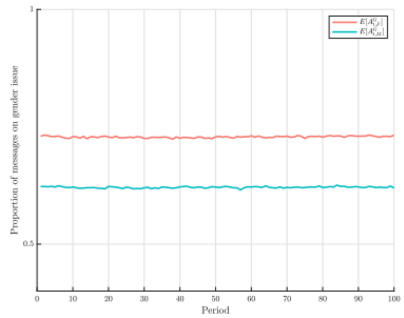
3. Simulations of Model Dynamic

To emphasize the fact that the asymmetry in feedback for messages on gender issues can be a sufficient cause for the emergence of an asymmetry in issue attention between politicians *F* and *M*, Figure SI1.1a presents the results of simulations obtained by assuming that the initial valuations of the two issues are the same for the two politicians (set to 0) but there is a sizeable gender issue feedback advantage for the *F* politician $\Delta\mu_F = 1 > \Delta\mu_M = .5$. We also assumed $r = 1$, $\gamma = .1$, $\sigma = 1$. Initially, both politicians devote the same attention to gender issues, but as the number of periods grows, an asymmetry in issue attention emerges. In Figure SI1.1b, we consider the case in which the initial valuations of the options correspond to the means of the feedback distributions. This amounts to assuming that agents *F* and *M* know the existence and the strength of the gender issue feedback advantage. In this case, there is a difference in issue attention between *F* and *M* from the start and it does not change (on average). Finally, in Figure SI1.1,c we consider the case in which the initial valuation of the gender issue by politician *F* is larger than the mean of the feedback distribution. In other words, *F* initially overestimates the gender issue feedback advantage. Accordingly the difference in attention to gender issues between *F* and *M* is initially large and it does down as *F* responds to feedback and adjusts her evaluation of the gender issue downward.

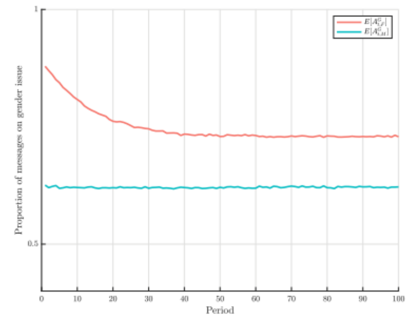
Figure SI1.1: Simulations of the dynamic of issue attention for *F* and *M*.



(a) Same initial valuations for politicians F and M, for tweets on both issues. $V_{F,1}^{GI} = V_{M,1}^{GI} = V_{F,1}^{other} = V_{M,1}^{other} = 0$



(b) Initially correct valuations for politicians F and M (equal to the means of the feedback distributions). $V_{F,1}^{GI} = \mu_F^{GI}$, $V_{M,1}^{GI} = \mu_M^{GI}$, $V_{F,1}^{other} = \mu_F^{other}$, $V_{M,1}^{other} = \mu_M^{other}$



(c) F initially overestimates the gender issue feedback advantage, M has correct initial valuations of the two issues. $V_{F,1}^{GI} = 2\mu_F^{GI}$, $V_{M,1}^{GI} = \mu_M^{GI}$, $V_{F,1}^{other} = \mu_F^{other}$, $V_{M,1}^{other} = \mu_M^{other}$

Note: Figure based on 100,000 simulations with $\mu_F^{GI} = 1$, $\mu_M^{GI} = .5$, $\mu_F^{other} = \mu_M^{other} = 0$, $r_F = r_M = 1$, $\gamma = .1$ and $\sigma = 1$.

B. Deep Learning Tweet Issue Classifier

This section describes how we used the BERT language model to classify tweets and predict feedback.

1. Why using BERT to classify tweets?

BERT-based text classifiers offer three advantages over other machine learning classifiers. First, they perform better than 'bag-of-words' classifiers which are most often used in the social sciences (Grimmer and Stewart, 2013). By contrast to the latter, BERT is sensitive not only to word frequencies or word sequences but also to context effects. The mathematical representation of a word depends on the other words that come before and after in the text. BERT performs so well because of this sensitivity to bi-directional dependency in word meaning. Second, BERT is pre-trained on a vast amount of data (the text of all Wikipedia articles) to learn a rich language representation but can then be 'fine-tuned' for specific tasks such as classification. Most text classifiers based on machine-learning techniques are trained from scratch on a particular dataset. If the data is of limited size, performance suffers. Classifiers that are pre-trained on large amounts of text but cannot be fine-tuned are limited by the fact that word representations are not adapted to the particular task at hands (in our case, identifying tweets on gender issues). Our BERT-based model overcomes the limitations of these two earlier approaches. Third, there exists a multi-lingual version of BERT that can be used with text written in more than 100 languages. This implies that it is not necessary to translate the texts before inputting them into the model. This was vital for us, as Spanish politicians regularly tweet in Spanish (Castilian), Basque, Catalan and Galician.

2. Human Coding Stage

Supervised machine learning algorithms require a set of tweets which are correctly labeled as being on gender issues or not. We manually classified tweets as follows. First, we developed coding guidelines by creating a list of issues related to gender. Second, we selected a random sample of 19,377 tweets from that topic to be the training set and another 1975 tweets as a test set. To maximize the information contained in the training set, we over-sampled tweets on gender issues using an unsupervised topic model (LDA). We sampled tweets from one of the topics constructed by the model which contained many of words related to gender issues. The test set was sampled without over-sampling, to be representative of the whole sample. Third, we trained research assistants to code tweets independently, and resolved inter-coder disagreement or ambiguous cases by discussing with them the tweets on which such disagreement occurred. Based on a pilot study, we decided that each tweet was to be coded by two research assistants and in case of disagreement, we would search for a consensus solution. They reached an inter-coder reliability of 0.89 measured as Fleiss' Kappa which is considered a very high agreement (Landis and Koch, 1977). Disagreement occurred in only 5.2% of cases.

3. Fine-tuning BERT-based artificial neural network models

To fine-tune the algorithm, we use a 10-fold cross validation (for an introduction see Hastie, Tibshirani and Friedman, 2009). This was implemented with Python relying on the Pytorch machine learning library by adapting publicly available code provided as part of the 'Transformers' library of language models (Wolf et al., 2019), available at https://github.com/huggingface/transformers. We created our main script by editing the provided 'run glue.py'. We used all the default training parameters except for the following parameters which we found would lead to higher performance on the kind of data we are using: per gpu train batch _size=64, learning _rate= 2e-5, warmup _steps=.1, max grad _norm=1.0, num train _epochs=1.0. The model was trained using a distributed training procedure on a GPU equipped workstation configured to perform fp16 computations (NVidia RTX 3090).

4. Accuracy of Classification

Our model achieved an excellent classification performance. More precisely, it obtained a precision of .90 and a recall of .79. This means that 90% of tweets the model classified as being on gender issues are actually on gender issues and that 79% of gender issue tweets are classified as being on gender issues. For comparison, we also trained a model that adopts the ‘bag-of-words’ approach, the na¨ıve Bayes classifier.[1] Our fine-tuned BERT classifier produces about one third of the mistakes produced by the na¨ıve Bayes classifier (39 vs. 140). Table SI2.1 reports the confusion matrices for the predictions of our fine-tuned model and of the naive Bayes classifier on the test data.

To develop an intuition for the quality of the model predictions, we computed the coefficient of inter-rater reliability (Fleiss’ kappa) by assuming the fine-tuned BERT model is a rater, and human categorization by the research assistants is another rater. The obtained coefficient is .83, which is a level generally considered as ‘almost perfect agreement.’ The same coefficient for the naïve Bayes classifier is .55, which is generally considered as ‘moderate agreement.’

Table SI2.1: Confusion matrices for the BERT gender issue classifier and the naïve Bayes gender issue classifier on the validation dataset (N=1974).

		Model Prediction		
		No	Yes	Total
Human Coding	No	1832	11	1843
	Yes	28	103	131
	Total	1860	114	

(a) BERT Multilingual Cased Classifier

		Model Prediction		
		No	Yes	Total
Human Coding	No	1735	108	1843
	Yes	32	99	131
	Total	1767	207	

(b) Naïve Bayes Classifier

[1] We use the Multinomial Naive Bayes model of the scikit-learn machine learning Python package. For details see: https://scikit-learn.org/stable/modules/naive_bayes.html.

5. BERT-based regression model for tweet feedback prediction

To predict feedback, we relied on an artificial neural network based on BERT Multilingual-cased. Model training is performed following similar steps as in the model identifying tweets on gender issue, but the output layer in this case is not a classification layer, but a linear regression layer which takes as an input the 768 dimension vector output by BERT and outputs predicted feedback as a linear combination of the vector elements.

We split our dataset of tweets into two sets of approximately the same size, by creating a random split of politicians such that all the tweets published by a given politician would fall in one of the two sets (call them set A and set B). We adopted this politician-level split of the data to prevent the algorithm from learning about the communication style of individual politicians and the popularity of the topic of gender issues among their followers - which it could theoretically do even though no explicit pointers to politicians are part of the input data.

We constructed the measure of feedback for a given tweet by starting with the number of retweets, taking out the politician level time trend (eq. 6), taking out day fixed effects, and then normalizing within politicians. Unlike the measure used in the main analyses ($FB_{i,m}$), this measure includes day fixed effects. This step was not necessary when constructing the main measure because day fixed effects could be included in the regression analyses. Yet, such *post-hoc* inclusion of fixed effects is not possible in this case because we aim to use the trained model for out-of-sample predictions and thus need to remove the effects of day to day variations at the training stage.
We used all the tweets in set A to fine-tune the model and applied the resulting model to predict the success of tweets in set B. We then used the tweets in set B to fine-tune the model and applied the trained model to predict feedback for the tweets in set A. This procedure allowed us to produce out-of-sample predictions of the amount of feedback expected by a tweet, just based on its semantic content. We would like to emphasize that no features of the tweet author were included as inputs, just the text of the tweet. The correlation between out-of-sample prediction and true feedback was about .50 (.52 for the model fine-tuned on set A and .49 for the set B model).

C. Descriptive Statistics

Table SI3.3: Summary Statistics: Tweets

Statistic	N	Min	Median	Mean	Max	St. Dev.
Tweet is on gender issues	1,583,917	0	0	0.06	1	0.24
Tweet is on Catalan independence	1,583,917	0	0	0.09	1	0.29
Writer is female politician	1,583,917	0	0	0.38	1	0.49
Writer is left-wing politician	1,583,917	0	1	0.56	1	0.50
Number of retweets	1,583,917	0	3	56.37	42,244	385.18
Number of likes	1,583,917	0	5	96.50	70,085	709.03
Number of replies	1,583,917	0	0	9.75	25,633	82.85
Feedback measure (based on retweets)	1,583,917	−6.81	−0.11	−0.00	14.99	1.00
Share of female retweeters	678,859	0.00	0.38	0.39	1.00	0.25
Thread length	1,583,917	1	1	1.00	5	0.02
Tokens	1,583,917	3	21	22.34	95	11.84
Hashtags	1,583,917	0	0	0.45	30	0.92
Mentions	1,583,917	0	1	1.09	50	1.84
Emojis	1,583,917	0	0	0.28	140	1.07
Contains picture	1,583,917	0	0	0.31	1	0.46
Contains link	1,583,917	0	0	0.40	1	0.49
Sentiment score	1,582,931	0.00	0.19	0.27	1.00	0.26

Note: Tokens are words or other symbols (mentions, emojis, etc.). Mentions are references to other Twitter users. The share of female retweeters is only calculated for tweets starting from 2018 with at least one identified retweeter. The sentiment score could not be computed for approximately 1000 tweets.

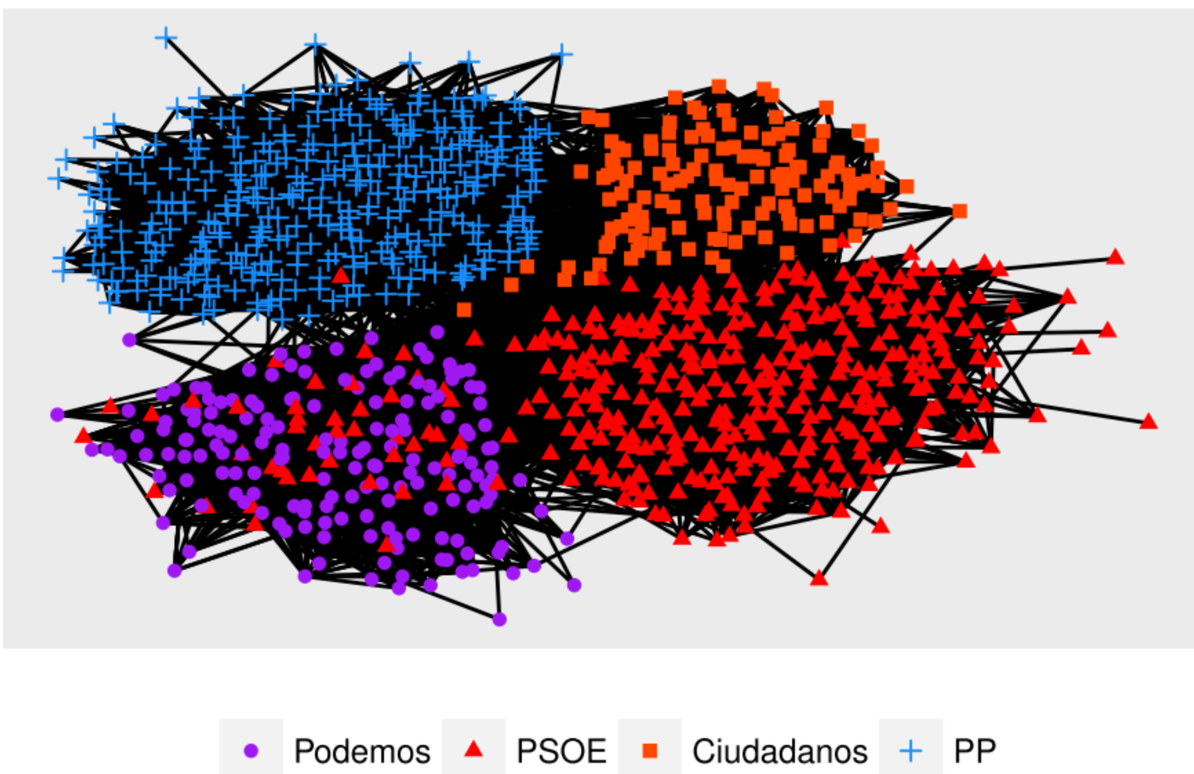
Table SI3.4: Summary Statistics: Politicians

Statistic	N	Min	Median	Mean	Max	St. Dev.
Share of tweets written on gender issues	1,265	0.00	0.04	0.07	0.60	0.08
Share of tweets written on Catalan independence	1,265	0.00	0.03	0.07	0.85	0.11
Female	1,265	0	0	0.44	1	0.50
Left-wing	1,265	0	1	0.57	1	0.50
Followers	1,184	74	3,169	23,699.10	2,390,647	120,038.80
Following	1,184	7	1,133	1,756.27	98,465	3,691.07
Tweets written since joining Twitter	1,184	133	8,227	13,168.80	134,620	15,156.80
Tweets written in sample period	1,265	33	652	1,252.11	29,172	2,059.92
Average number of retweets	1,265	0.13	6.87	35.54	2,651.50	136.06
Standard deviation of retweets	1,265	0.37	11.45	76.72	3,660.48	234.62
Average number of likes	1,265	0.32	9.41	62.00	5,040.03	274.30
Standard deviation of likes	1,265	0.82	15.45	131.20	6,687.89	456.71
Average number of replies	1,265	0.00	0.63	5.68	575.24	27.60
Standard deviation of replies	1,265	0.00	1.77	17.16	2,654.09	88.27
Average number of tokens	1,265	7.36	22.63	23.04	45.14	5.47
Average share of female retweeters	1,262	0.00	0.38	0.38	0.83	0.12

D. Evidence for Retweets as Positive Feedback

Figure SI4.3 plots the network of retweets between Members of Parliament of Spain's four major parties (n=527). Members of Parliaments from one of the smaller parties were excluded to facilitate visualization. Vertices correspond to politicians. An edge exists if one politician retweeted another politician in our sampling period or vice versa. The figure shows that most retweets happen within parties. We interpret this as evidence that retweets are used as positive feedback.

Figure SI4.3: Retweeting Network between Politicians. Each vertex is politician. An edge exists between two vertices of the two politicians retweeted at least one tweet by the other politician.



Note: The colors correspond to elected politicians of Spain's four major parties (Podemos, PSOE, Ciudadanos, PP) during our sample period (2016-2019).

E. Gender Issue feedback Advantage – Robustness

1. Differences Between Left- and Right-Wing Parties

Model 4 in Table 3 in the main text tests if the gender issue feedback advantage is driven by one side of the political spectrum. We could conjecture that left-leaning politicians might receive more positive feedback for addressing gender issues or that a stronger feedback advantage for female politicians might be more pronounced among right-leaning politicians.[2] However, when we interact our variables of interest ($GI_{i,m}$, M_i) with a dummy equaling 1 if politician i belongs to a left-leaning party L_i, we do not find that our effects depend on the politician's ideological leaning. We coded parties as follows:

Left-leaning parties: ASG, AVANCEM, Bildu, BNG, CHA, Coalici´on Caballas, COMPROMIS, CpM, CUP, EM, ERC, Eusko Alkartasuna, GENTxFORMENTERA+PARTIT SOCIALISTA DE LES ILLES BALEARS, Geroa Bai, ICV, INDEPENDENT, MDyC, MES PER MALLORCA-PSM-´ ENTESA-INICIATIVAVERDS, MES PER MENORCA, NCa, Podemos, PRC, PSOE, UPL´

Right-leaning parties: CCa-PNC, Ciudadanos, EAJ-PNV, EL PI-PROPOSTA PER LES ILLES, Foro Asturias, JxCat, PAR, PDECAT, PP, PPL, UPN, VOX

2. Feedback Measure Based on Retweets, No Within-Politician Normalization

The within-politician normalization stage of the construction of our feedback measure presumes that the psychologically relevant feedback for a politician is the amount of retweets *relative to a politician-specific baseline*. Even though work on adaptive aspirations suggests that this assumption is realistic, it is possible that politicians who tend to receive more retweets pay more attention to retweets than those who tend to receive few retweets. This possibility is assumed away by the within-politician normalization step in the construction of the

[2] Note that in the case of Spain, a left-right classification also firmly aligns with a divide between progressive and socially conservative parties (Rama et al., 2021).

feedback measure. Relatedly, within-politician normalization implies that the impact of one more retweet on feedback will differ between politicians, such that one more retweet has a larger impact on the feedback measure for politicians who generally receive few retweets. Because female politicians receive fewer retweets than male politicians, this makes the comparison between politicians of these two groups tricky, possibly leading to an inflated estimate of the difference in feedback received by female and male politicians.

To address these possibilities, we replicate our baseline specification (Model 3 in Table 3) removing the within-politician normalization step in the construction of the feedback measure. We find that results are similar to our main results (see Model 2 in Table SI5.5). The difference between female and male politicians is statistically significant ($p < 0.01$).

3. Feedback Measure Based on the Number of 'Likes'

Model 3 in Table SI5.5 replicates Model 3 in Table 3 in the main text using likes instead of retweets to construct our feedback measure. In line with the main results, female politicians have a larger gender issue feedback advantage than male politicians. The difference is statistically significant ($p < 0.01$).

4. Feedback Measure Based on Number of 'Replies'

Replies are a third possible source of feedback on Twitter. But by contrast to retweets, replies can be positive or negative feedback. Since they often originate from political opponents (Conover, Ratkiewicz and Francisco, 2011), it does not come as a surprise that most but this correlation is much smaller than the correlation between the like-based and the retweet-based feedback measures ($\rho = .85$).

Taken together, these three analyses cast doubt that tweets with a relatively large number of replies were received more positively by the public than other tweets. A large number of replies can also indicate that a tweet was received critically.

Despite this caveat, we replicated Model 3 of Table 3, with the reply-based measure of feedback as the dependent variable instead of the retweet-based measure. We find that tweets on gender issues receive more replies but less so for male politicians. The difference between female and male politicians is statistically significant ($p < 0.01$, see Model 4 in Table SI5.5). However, if we include the retweet-based feedback measure in the regression as an additional control, the coefficient for 'GI' becomes negative and the difference between female and male politicians is not significant anymore ($p > 0.1$, see Model 5 in Table SI5.5). This result means that controlling for the popularity of a tweet, gender issue tweets receive relatively few replies. We find a similar pattern if we use the like-based feedback measure as a control. Thus, we do not find clear evidence that tweets receive more replies because they are on gender issues.

Table SI5.5: Additional Robustness Checks for Table 3: Female politicians receive more positive feedback for tweeting on gender issues.

Dependent Variables:	Retweets	Non-normalized retweets	Likes	Replies	Replies
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Gender issues	0.28**	0.34**	0.26**	0.08**	-0.06**
	(0.02)	(0.03)	(0.02)	(0.02)	(0.01)
Gender issues * Male politician	-0.12**	-0.14**	-0.10**	-0.07**	-0.00
	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)
Part of thread	0.29†	0.42*	0.25	0.24**	0.10
	(0.15)	(0.17)	(0.16)	(0.09)	(0.12)
Tweets on day by politician	-0.01**	-0.01**	-0.01**	-0.00**	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Retweets					0.49**
					(0.01)
Fixed-effects					
Politician	Yes	Yes	Yes	Yes	Yes
Day	Yes	Yes	Yes	Yes	Yes
Hour of day	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Squared Correlation	0.02	0.52	0.02	0.01	0.24
Observations	1,583,917	1,583,917	1,583,917	1,583,308	1,583,308

Note: Linear regressions of tweet feedback on politicians' gender and issue of the tweet (variations of equation 7). Model 1 repeats our Model 3 in Table 3, our baseline specification. Model 2 uses non-normalized retweets as dependent variable. Model 3 uses likes. Model 4 uses replies. Finally, Model 5 uses likes as dependent variable while controlling for retweets. Standard errors as reported in parentheses are clustered by politician †p<0.1; *p<0.05; **p<0.01

F. Responsiveness to Feedback – Robustness

This section extends the discussion about the robustness checks of the responsiveness models discussed in the body of the paper. The model estimation results are reported in Table SI6.6.

1. Details on the Additional Specifications Reported in Table 4 in the body of the paper

We first account for individual trajectories in politicians' attention to gender issues over time. This is relevant since female politicians increase their attention to gender issues more than male politicians during our study period. As explained in Section 4.2, the feedback measure (on which issue valuations are based) already includes a politician-specific trend. This makes valuations more comparable over periods even when politicians are on different time trends. We do the same for issue attention by including a linear time trend for each politician (Table 4, Model 4). The coefficient for the trend is highly significant and increases the fit of the model. Yet, the estimated coefficients for the issue valuations hardly change. This is noteworthy as the trend is arguably endogenous to feedback: politicians who consistently receive more positive feedback for tweeting on gender issues will be on a steeper trend.

Next, we want to dispel concerns that serial correlation might bias our estimates. We include the lagged dependent variable (the share of tweets written on gender issues in the last month) as a control (Table 4, Model 5). The additional variable has a large positive coefficient and the model fit increases, but the estimated coefficients for the issue valuations remain similar to Model 3 in Table 4.

Finally, we address the possibility that issue attention is influenced by peer effects. Even though the month fixed effect already captures common patterns in issue attention that affect all politicians equally, it could be that politicians are more strongly affected by the behavior of politicians from the same gender. In Model 6 in Table 4, we include the average attention to gender issues by politicians of the same gender (male or female) as a control. The estimated coefficient is imprecisely estimated.

This suggests that this sort of peer effects does not play an important role.

2. Feedback Measure Based on 'Likes'

We chose retweets over likes to construct our feedback measure because it allowed us to learn about the feedback givers' gender. Still, our theory of reinforcement learning should also apply to likes as a form of feedback. Likes have the advantage that they unambiguously stand for positive feedback. Hence, if our responsiveness results replicate using likes instead of retweets, it provides further evidence that politicians are responsive to positive feedback.

The results are reported in Table SI6.6, Model 1. They are similar to what we obtained with the feedback measure based on retweets. Responsiveness coefficients all have the same sign as in the main results and are significant. Point estimates are somewhat attenuated, but close. Again, the difference in responsiveness between female and male politicians is not statistically significant (p-value>0.1). We conclude from this analysis that politicians are also responsive to likes as a form of feedback.

3. Feedback Measure Based on 'Replies'

As discussed in Appendix F, due to the sometimes negative nature of replies and uncertain expectations regarding how politicians respond to negative feedback, we did not necessarily expect that we could replicate our results using the number of replies instead of retweets to construct the feedback measure. Nevertheless, Model 2 of Table SI6.6 shows that attention to gender issues correlates positively with the reply-based feedback measure. Effect sizes are somewhat smaller but comparable to those obtained in the baseline analyses (Table 4, Model 3).

This pattern probably results from a positive correlation between number of replies and number of retweets. Therefore, we replicate the analysis by including both the previously introduced retweetbased feedback measure and the reply-based feedback measure of feedback. As can be seen in Model 3 of Table SI6.6, the coefficient on the retweet-based feedback measure is almost the same as in the baseline analyses (Table 4, Model 3) whereas the effect of replies disappears. This reveals that the reply-based feedback does not affect issue attention in a consistent way. Note that we are *not* claiming that replies do not have an effect. Instead, it is likely that the absence of statistical effect reflects unobserved heterogeneity. For example, depending on the personality of the politician and the tone of the reply, replies might either increase or decrease issue attention. We leave further investigation of replies to future research.

4. Feedback Measure Based on Retweets, No Within-Politician Normalization

The results are reported in Table SI6.6, Model 4. They are similar to the baseline results reported in the body of the paper. Both female and male politicians are generally responsive to feedback. Effects are statistically significant but there is no significant difference in responsiveness between the groups in any of the models (p-value>0.1).

5. Alternative Time Period Used to Compute Issue Attention

To show that our main results do not depend on the particular choice of time period for computing issue attention (months), we replicate the specification using weeks instead of months. Model 5 in Table SI6.6 shows that our results hold. Again, both female and male politicians are generally responsive to feedback but no difference between the two social categories can be detected. A higher valuation of gender issues increases the attention to the issue whereas a higher valuation of other issues can lead to a crowding out.

6. Alternative Regression Weights

To avoid concerns that our main results could be driven by the specific weighting scheme we used in the model estimations, we replicate our analyses by weighting each politician month cell equally, independently of the actual number of tweets written in the politician-month cell. Our main result holds, yet, there are some differences. For female politicians, the estimated responsiveness parameters remain stable. Model 6 in Table SI6.6 shows that a higher valuation of gender issues is associated with a higher share of tweets written on the issue among female politicians. For male politicians, the effects is attenuated and only barely significant ($p < 0.1$). We believe that this makes sense, considering that many male politicians write few tweets on gender issues and we need to focus on the set of politician-month cells containing more tweets to find significant effects. Giving equal weight to cells with too few underlying tweets creates too much noise. Nevertheless, the difference in effect sizes is not statistically significant in any specification.

7. Alternative Reference Category

In our analyses, we assumed that a politician who chooses the issue of a tweet selects between 'gender' and 'other' based on her valuation of the two issues. Yet, it is not psychologically realistic that politicians have a clear mental representation of the valuation the 'other' issue. To address this potential limitation, we coded tweets for another substantive category, that of 'Catalan independence' and used it as an alternative reference category.[3] Model 7 in Table SI6.6 shows that using the alternative reference category does not affect the main interpretation of our results. If politicians receive more positive feedback for addressing the topic of gender issue, they tend to write more about it. On the other hand, positive feedback for tweets on Catalan independence have a slightly negative (though not significant) effect. Comparing it to the effect of the 'other' category V_{other}' reveals that the effect is relatively small. This can be explained by the fact that two specific issues compete less about attention compared to one specific issue competing with all other issues at once.

[3] We used an approach similar to that adopted to code the 'gender issue'. We hired research assistants to code about 12,000 tweets. We used about 10,000 tweets for training a BERT classifier, about 2,000 tweets for model validation and applied the model on the remaining tweets.

8. Placebo Test

To implement the placebo test, we randomly swap politicians' issue valuations with the feedback-based valuation of another politician of the same gender (male or female). The results are reported in Model 8 in Table SI6.6. The coefficients of issue valuations are close to zeros and do not reach statistical significance.

Table SI6.6: Robustness Checks for Table 4: Responsiveness to feedback of male and female politicians

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Valuation}^{\text{gender issues}}_{\text{Female politician}}$ likes	0.12^{**}							
	(0.04)							
$\text{Valuation}^{\text{gender issues}}_{\text{Male politician}}$ likes	0.11^{**}							
	(0.04)							
$\text{Valuation}^{other}_{\text{Female politician}}$ likes	$-0.06^{\dagger}$							
	(0.03)							
$\text{Valuation}^{other}_{\text{Male politician}}$ likes	-0.04							
	(0.03)							
$\text{Valuation}^{\text{gender issues}}_{\text{Female politician}}$ replies		0.10^{*}	0.01					
		(0.04)	(0.02)					
$\text{Valuation}^{\text{gender issues}}_{\text{Male politician}}$ replies		0.10^{**}	0.00					
		(0.03)	(0.02)					
$\text{Valuation}^{other}_{\text{Female politician}}$ replies		-0.05	-0.01					
		(0.03)	(0.02)					
$\text{Valuation}^{other}_{\text{Male politician}}$ replies		-0.09^{**}	$-0.04^{\dagger}$					
		(0.03)	(0.02)					
$\text{Valuation}^{\text{gender issues}}_{\text{Female politician}}$ retweets			0.14^{**}	0.12^{**}	0.15^{**}	0.19^{**}	0.16^{**}	0.00
			(0.04)	(0.03)	(0.03)	(0.05)	(0.05)	(0.02)
$\text{Valuation}^{\text{gender issues}}_{\text{Male politician}}$ retweets			0.12^{**}	0.11^{**}	0.10^{**}	$0.07^{\dagger}$	0.14^{**}	0.00
			(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.02)
$\text{Valuation}^{other}_{\text{Female politician}}$ retweets			-0.07^{*}	-0.08^{**}	-0.12^{**}	-0.04		0.01
			(0.03)	(0.03)	(0.02)	(0.05)		(0.03)
$\text{Valuation}^{other}_{\text{Male politician}}$ retweets			-0.00	-0.03	-0.05^{*}	-0.05		0.02
			(0.03)	(0.03)	(0.02)	(0.04)		(0.03)
$\text{Valuation}^{\text{Catalan independence}}_{\text{Female politician}}$ retweets							-0.05	
							(0.05)	
$\text{Valuation}^{\text{Catalan independence}}_{\text{Male politician}}$ retweets							-0.03	
							(0.04)	
$\hat{\gamma}$ likes	0.07							
$\hat{\gamma}$ retweets			0.07	0.07	0.15	0.08	0.06	0.07
$\hat{\gamma}$ replies		0.1	0.3					
Fixed-effects								
Politician	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes		Yes	Yes	Yes

Week					Yes			
Fit statistics								
Squared Correlation	0.57	0.57	0.57	0.57	0.36	0.59	0.63	0.57
Observations	18,482	18,469	18,469	18,482	74,599	18,482	16,626	18,482

Note: Logistic regression of the share of tweets written on gender issues on the valuation of gender issues and other issues (see model in equation 8). Model (1) uses likes instead of retweets to estimate issue valuation. Model (2) uses replies. Model (3) uses replies while controlling for the retweets-based valuation. Model (4) uses an alternative feedback measurement (non-normalized). Model (5) uses weeks instead of months as time frame. Model (6) weights each month equally (instead of each tweet). Model (7) uses an alternative reference category (Catalan independence instead of 'other issues'). Model (8) is a placebo test with randomly swapped valuations. $\hat{\gamma}$ is the estimated updating parameter used to calculate issue valuations. All regressions use cell-size regression weights, i.e., number of tweets published by politician i in period p (N_{ip}). Standard errors are clustered by politician: $^{\dagger}p<0.1$; $^{*}p<0.05$; $^{**}p<0.01$.

G. Mechanisms – Additional results

Table SI7.7: Mechanisms for the difference in gender issue feedback advantage between female and male politicians - Regression results complementing Figure 2

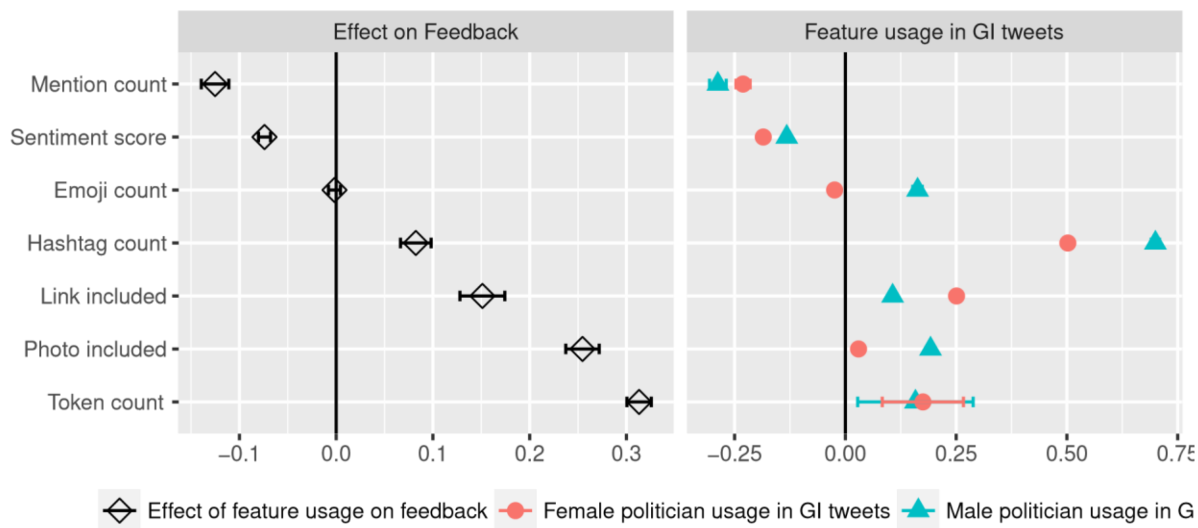
Dependent Variable:	Retweets	Predicted Feedback	Retweets	Retweets	Retweets	Retweets female users	Retweets male users
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gender issues	0.28**	0.18**	0.09**	0.15**	0.27**	0.30**	0.17**
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Gender issues * Male Politician	-0.12**	-0.02	-0.11**	-0.10**			
	(0.02)	(0.01)	(0.02)	(0.02)			
Tweet quality			1.08**				
			(0.01)				
Sentiment score				-0.07**			
				(0.00)			
Token count				0.31**			
				(0.01)			
Hashtag count				0.08**			
				(0.01)			
Mention count				-0.12**			
				(0.01)			
Emoji count				-0.00			
				(0.00)			
Photo/video included				0.25**			
				(0.01)			
Link included				0.15**			
				(0.01)			
Tweets on day by politician	-0.01**	-0.00**	-0.00**	-0.01**	-0.01**	-0.01**	-0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Part of thread	0.29†	0.30**	-0.04	0.00†	0.34*	0.38**	0.31*
	(0.15)	(0.08)	(0.09)	(0.00)	(0.16)	(0.15)	(0.15)
Fixed-effects							
Politician	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour of day	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Squared Correlation	0.02	0.20	0.31	0.19	0.06	0.02	0.02
Observations	1,583,917	1,583,917	1,583,917	1,582,931	1,105,698	1,105,554	1,105,698

Note: Linear regressions of tweet feedback on politicians' gender and issue of the tweet (variations of equation 7). Model 5-7 only consider tweets starting from 2018 with retweeter information. Standard errors are clustered by politician. †p<0.1; *p<0.05; **p<0.01.

1 Tweet Style

Figure SI7.4 indicates that, compared to male politicians, female politicians do not systematically use more features that attract positive feedback in their gender issue tweets. More specifically, the left panel shows the effect of different features on standardized feedback. The right panel shows the average usage of those features in gender issue tweets (compared to other tweets) separately for female and male politicians. However, female politicians do not use features that systematically attract more feedback when tweeting on gender issues, relative to male politicians. Hence, style differences are unlikely to contribute much to the difference in gender issue feedback advantage.

Figure SI7.4: Female politicians do not use Tweet features that attract more positive feedback when tweeting on gender issues



Note: Left panel: Each diamond represents the estimated effect of including a give feature in a Tweet on feedback. The effects were obtained from a linear regression of feedback on the vector of features for each tweet. Right panel: Each point represents usage of given feature in gender issue tweets relative to other tweets, separately for female and male politicians. It shows that female politicians do not use popular features (features towards the lower part of the figure) more often in their gender issue tweets. Bars represent 95% confidence intervals (sometimes not visible, as they are close to zero).

2 Coding the Gender of Retweeters

We infer the gender of Twitter users based on the Twitter username. For this, we use Genderize.io, a commercial online service that predicts if a name is male or female. Twitter only allows to retrospectively download information about up to 100 retweeters per tweet. Furthermore, some of their Twitter usernames were not indicative if the retweeter was male or female. Still, for the average tweet in our sample, we obtained a classification for 61% of the retweeters. We estimate the absolute number of female or male retweeters by multiplying the absolute number of retweeters with the estimated share of female and male retweeters of each tweet. Finally, we apply the same steps of feedback normalization (see Section 4.2) to the retweets of female and male retweets. We used this for Figure 3 and Models 6 and 7 in Table SI7.7.