

Weeks, Ana Catalano; Meguid, Bonnie M.; Kittilson, Miki Caul; Coffé, Hilde

Working Paper

Response to "Do Radical-Right Parties Use Descriptive Representation Strategically? A Replication of Weeks et al. (2023)"

I4R Discussion Paper Series, No. 150

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Weeks, Ana Catalano; Meguid, Bonnie M.; Kittilson, Miki Caul; Coffé, Hilde (2024) : Response to "Do Radical-Right Parties Use Descriptive Representation Strategically? A Replication of Weeks et al. (2023)", I4R Discussion Paper Series, No. 150, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/302284>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



No. 150

DISCUSSION PAPER SERIES

Response to “Do Radical-Right Parties Use Descriptive Representation Strategically? A Replication of Weeks et al. (2023)”

Ana Catalano Weeks

Bonnie M. Meguid

Miki Caul Kittilson

Hilde Coffé

This paper responds to:

Guinaudeau, B., M. Jankowski. 2024. Do Radical-Right Parties Use Descriptive Representation Strategically? A Replication of Weeks et al. (2023). *I4R Discussion Paper Series* No. 149. Institute for Replication.

September 2024

I4R DISCUSSION PAPER SERIES

I4R DP No. 150

Response to “Do Radical-Right Parties Use Descriptive Representation Strategically? A Replication of Weeks et al. (2023)”

**Ana Catalano Weeks¹, Bonnie M. Meguid², Miki Caul Kittilson³,
Hilde Coffé¹**

¹University of Bath/Great Britain

²University of Rochester/USA

³Arizona State University, Tempe/USA

SEPTEMBER 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research

Hohenzollernstraße 1-3
45128 Essen/Germany

www.i4replication.org

Response to “Do Radical-Right Parties Use Descriptive Representation Strategically? A Replication of Weeks et al. (2023)”

Ana Catalano Weeks¹, Bonnie M. Meguid², Miki Caul Kittilson³, Hilde Coffé⁴

Abstract:

Guinaudeau and Jankowski reassess our recent study on the use of *strategic descriptive representation* among political parties in Europe. The authors successfully replicate the vast majority of our findings and perform a number of additional robustness checks. They claim that one of our key findings is sensitive to the inclusion of one observation (the Front National, FN, 2012), and that alternative measurement or modeling strategies return different results. In this response, we address each claim in turn. We apply influential case diagnostics to detect all influential cases in our multilevel models, so as not to arbitrarily delete one influential observation but not another. On removing *all* influential cases, our results remain substantially the same. More importantly, because we do not agree with arbitrarily dropping observations, our findings are robust to different handling techniques for influential cases in multilevel models which downweight influential cases. Further, and in line with our original mixed methods approach, we provide an additional influential case study of the use of strategic descriptive representation by the FN in 2012, which is supportive of our theory and quantitative evidence. Finally, we respond to questions about our measurement and modeling decisions by highlighting the theoretical framework and scholarly literature that informs these decisions, which is largely disregarded by GJ.

¹ University of Bath, Dept of Politics, Languages & International Studies, 1 West North, Claverton Down, Bath BA2 7AY, United Kingdom, a.c.weeks@bath.ac.uk

² University of Rochester, Dept of Political Science, 306 Harkness Hall, Rochester, NY 14627, USA, bonnie.meguid@rochester.edu

³ Arizona State University, School of Politics and Global Studies, 975 S. Myrtle Ave., Tempe, AZ, USA, miki.kittilson@asu.edu

⁴ University of Bath, Dept of Politics, Languages and International Studies, Building 1 West North 2.36, Bath BA2 7AY, United Kingdom, h.r.coffe@bath.ac.uk

Historically men-dominated radical right populist (RRP) parties are increasingly including women among their ranks. In an article published in the *American Political Science Review* (Weeks, Meguid, Kittilson and Coffé 2023), we build on the literature on party strategic behavior to develop a new theory explaining women's increased descriptive representation in RRP parties: RRP parties use *strategic descriptive representation*, electing women MPs when they lack women voters and are losing voter support. We test this argument using mixed methods, pairing quantitative analysis of European parties over time with illustrative case studies to assess the plausibility of our argument. Our results show consistent evidence that RRP parties use strategic descriptive representation in the conditions we describe, in both the quantitative analysis and qualitative analysis of real party campaigns in Switzerland and the Netherlands.

Guinaudeau and Jankowski (GJ) reassess our quantitative findings among RRP parties and argue that these results are influenced by the case of the Front National in 2012. Additionally, they claim that our findings are not robust to alternative modeling strategies (REML estimation for multilevel models or country fixed effects), and that measuring our main explanatory variable in a different way changes the findings.⁵ After seeing our initial response,⁶ GJ added a rejoinder questioning one of our approaches to identifying influential cases and providing a new analysis of gender and candidate selection in the FN 2012.

In the following response, we address each of GJ's claims in turn. First, as GJ note, computer reproducibility is nearly perfect. GJ are able to verify nearly all of our results using our data and code. There is one exception, which as GJ explains is related to an R package that we used to output tables. To create our regression tables, we utilized a popular R package called Stargazer. GJ claim that this package affects the p-values of "some models." We reviewed the results of all regression models in our article, and we find that the use of Stargazer only changes the output of *one* p-value in *one* table: In Table 1 Model 2, the p-value is reported as $p < 0.05$ when using Stargazer, when in fact it is $p=0.0548$. The issue does not affect any other variables in the analysis, and thus, does not change our conclusions.

Second, we respond to the critique about influential cases by carefully examining our mixed methods data in several ways. We first apply diagnostics for identifying influential cases in multilevel models which account for the nested structure of our data, which GJ fail to do. Best practices recommend identifying all influential cases (Ruiter

⁵ A first version of GJ's letter, then solo-authored by Jankowski and focusing exclusively on the influence analysis, was reviewed by the *American Political Science Review* and external reviewers as part of a Committee on Publication Ethics (COPE) process. His concerns were rejected, and the conclusion of this process was that no further action on our part was required.

⁶ I4R procedure allows the replicators to change their report after receiving a response from the authors. They then allow the authors to change their response as well.

and De Graaf 2010), so as to avoid arbitrarily deleting one influential observation but not another. When we systematically drop all influential cases (including those that clearly dampen the hypothesized effect, rather than only those that strengthen it), our key findings are robust. But removing – and thus ignoring – correctly-coded cases is not the recommended methodological approach. Instead, methodologists recommend employing models that explicitly recognize and downweight influential cases (Koller 2016; Van der Meer et al. 2010). When we employ these handling methods, including robust regression for multilevel models, again our main results hold.

Given that our data for RRP parties consists of a relatively small-N – which reflects the population of cases and which we transparently report – the case-based logic that we employ in our article is critical to strengthening the evidence for strategic descriptive representation. Our results should be evaluated in their entirety, rather than focusing only on the quantitative evidence as GJ do. To bolster this case-based analysis, here we provide an additional influential case study of the French Front National before and during the 2012 elections. We find that the party used strategic descriptive representation to promote women to winnable seats and in visible positions under exactly the conditions our argument predicts, in contrast to GJ's claim. In other words, this case provides important leverage for understanding party-strategic decisions about women's representation and should be studied rather than dropped.

Third, we respond to the additional modeling and measuring critiques. First, GJ prefer REML to MLE estimation for multilevel models. We show that this choice makes no difference to our findings. Second, GJ suggest the use of country fixed effects rather than multilevel models. We explain why that modeling choice is not appropriate for our nested data structure which includes country-level, time-invariant covariates, and we also perform Hausman tests which confirm the appropriateness of the use of multilevel models. Third, GJ propose an alternative way of measuring our main explanatory variable, suggesting that the share of men voters rather than male/female voter ratio be used. They misunderstand our ratio coding – which measures national-level share of men voting for the party divided by national-level share of women voting for the party – and so the alternative variable they suggest does not measure what they claim it does. Additional data collection would be necessary to test their hypothesis; nevertheless, the use of ratios is an appropriate and conventional way used in the literature to test our theory about gender gaps.

Fourth, in their rejoinder (“Comments on WMKC’s Response”), GJ miss the larger point we make: arbitrarily deleting correctly-coded observations is not appropriate, particularly given the relatively small sample size here. Instead, they critique our multilevel identification of influential cases. We show that their new estimation of Model 5 does not follow their own recommendations, and we provide a new analysis

that continues to show our results are substantially the same. Last, GJ offer a new analysis of gender and candidate selection within the FN. However, their additional statistical analyses do not test the central claims of our theory and introduce metrics that are not relevant to and do not challenge the importance of the FN 2012 case to the strategic descriptive representation phenomenon.

In summary, as we demonstrate in detail below, we disagree with GJ's critiques, and we provide strong evidence countering their claims. In what follows, we engage with their statistical focus for the purpose of directly responding to their points. But as we also highlight throughout, good social scientific research also is theory-driven, builds upon and advances a substantive literature, and reflects knowledge of the underlying cases. These are the goals that motivated our original article and continue to shape our approach in this response.

Identifying and Handling Influential Cases

We agree that it is important to investigate whether some observation or group of observations could have disproportionate influence in model estimation. In doing so, it is critical to properly account for the structure of the data being examined. Typical tools for measuring whether a model is influenced by one or a small set of observations assume linearity, which is not the case with the data in our models. We examine parties, which are nested within countries; in other words, the party-level observations are not independent, as is the assumption in linear models. This data structure guides our multilevel model choice in the original article. Specifically, party strategic choice sets are shaped by the actions of other parties in the national electoral system context. Therefore, we must employ diagnostics for influence in multilevel models, which allow for a dependence structure where observations belonging to the same higher level group can be correlated (see Ruiter and De Graaf 2010; Van der Meer et al. 2010).

Following these best methodological practices, we apply influential case diagnostics to detect all influential country cases in our multilevel models - the level at which these observations are nested. By applying the influential diagnostics as strictly as possible, we avoid arbitrarily deleting one influential observation but not another. We employ Cook's distance values, which provide a measure of the extent to which model parameters are influenced by (a set of) influential data on which the model is based. Figure 1 presents the Cook's distance values for all country level cases included in Model 5 of Table 1 (Weeks et al. 2023).

Figure 1: Diagnostics for Influential Cases in Multilevel Models: Cook's Distance for RRP parties

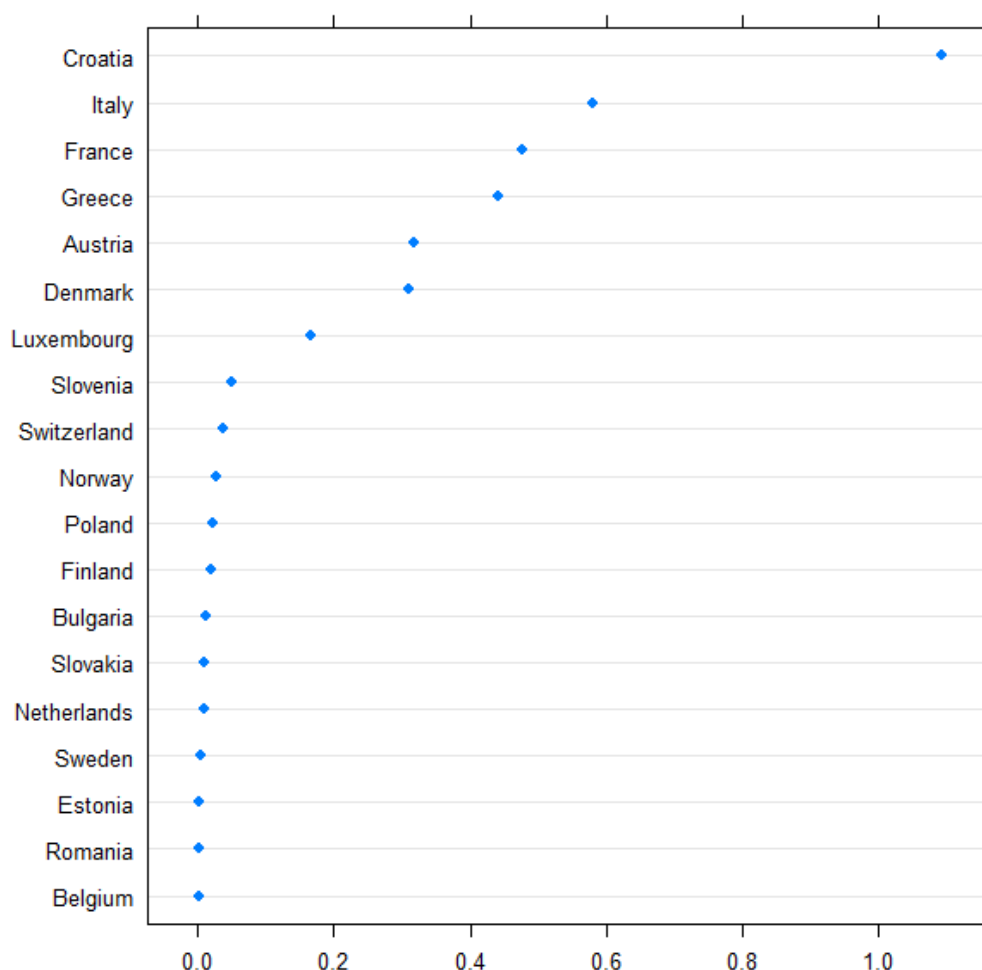


Figure 1 shows one clear outlier with Cook's distance greater than 1, which is not France but Croatia. However, several other countries also have relatively high values of Cook's distance. Using the standard cutoff criterion of $4/n$ (where n = number of cases; Belsley et al. 1980), the cutoff for influence is $4/19 = 0.2105$. According to this cutoff, six country cases are influential: Croatia, Italy, France, Greece, Austria, and Denmark.

Rather than single out one influential observation as was done by GJ, a strict reading of the influential diagnostics indicates that we should treat all six of these influential country cases for Model 5 the same. Not only does this ensure that objectivity is

maintained, but it also allows us to avoid arbitrarily deleting one influential case but not another. Indeed, following this methodology faithfully ensures that we do not just identify the cases that might reinforce the estimated relationship, but also the cases that weaken it (Ruiter and De Graaf 2010). Just focusing on one observation, such as France's Front National 2012, does not follow this systematic approach.

We replicated our results using the deletion process advocated by GJ, while additionally accounting for the multilevel data structure and excluding all influential cases, per the discussion above. As we show in Appendix A, contrary to the concerns raised by GJ, our findings are robust to the removal of influential cases. Our diagnostics suggest that while the case of the Front National in France strengthens our key association of interest, other influential cases clearly dampen the hypothesized effect. When *all* influential cases are removed, and not just the observation of FN in 2012, our key findings do not change. Appendix A also includes a response to GJ's rejoinder critique of our multilevel diagnostics, where we recalculate Cook's distance and drop influential cases based on the same exact model specification, using a modified version of our fullest model. Again, the results are substantially the same.

Handling Techniques for Influential Cases in Multilevel Models

While the results from deleting influential observations are supportive of our original conclusions, simply deleting *correctly coded* observations does not follow the recommended methodological approach.

The literature draws an important distinction between (and proper treatment of) observations that are influential and those that stand out because of possible coding errors. Recent best-practice recommendations for handling influential observations are clear about the inappropriateness of deleting correct observations. Aguinis et al. (2013) state that, "we recommend against the practice of automatically treating any data point as harmful... defining outliers in a negative way most often leads to simply removing these cases, a practice that may result in artificial range restriction (McNamara, Aime, & Vaaler, 2005)". Similarly, Van der Meer et al. (2010) conclude in their discussion of influential cases in multilevel modeling, "We do not advocate for deletion of higher level influential cases altogether in multilevel analysis" (p. 177). Koller (2016) also clearly maintains that, "It is important not to just remove and, thus, ignore outliers" (p. 13).

Deleting correctly-coded higher influence observations makes the sample an inaccurate representation of the population, and thus can lead to improper inferences about a population, particularly in a small population like the data on radical right populist

parties in Europe. Indeed, this search for statistical outliers and influential observations devoid of any context introduces the possibility that at some point, dropping observations without any substantive justification is simply manipulation of the sample to arrive at the desired outcomes. In our population, deleting the Front National, or even just the party in the 2012 election year when it achieved the second best electoral result in its lifetime, would make very little sense. The FN is one of the most electorally successful and oldest radical right populist parties in Europe, founded in 1972, and contesting national legislative and presidential elections since 1973 and 1974, respectively. To cite Bouvet (2013: 3), the “Front National prospered to become one of the most enduring, and arguably the most significant, European neo-populist party.” Indeed, many party experts, including Cas Mudde, have described the FN as the “prototype” of the radical right populist party family in Europe.⁷ Eliminating this party would be substantively misrepresenting the arc of the radical right populist party phenomenon in Europe.

Therefore, we turn to the approach recommended by methodologists (Koller 2016; Van der Meer et al. 2010) when facing influential cases: employing models which explicitly recognize and downweight influential cases. In Table 1, we summarize the results of robust regression models as applied to our puzzle of radical right women MPs. These approaches were identified based on their specific ability to handle multilevel models. As we see across the variety of specific modeling estimators/techniques, results consistent with our original *APSR* models emerge.

Table 1: Determinants of Women's Representation in Radical Right Populist Parties, Accounting for Influential Cases

	Model 1 (original Table 1, Model 5)	Model 2 (country fixed effect for France & adapted intercept) <i>Van der</i> <i>Meer et al.</i> 2010	Model 3 (country fixed effects for all influential countries & adapted intercept) <i>Van der</i> <i>Meer et al.</i> 2010	Model 4 (robust regression) <i>Koller 2016</i>
M/F Ratio _(t-1)	-0.605 (0.431)	-0.549 (0.418)	-0.769* (0.406)	-0.375 (0.516)
Vote Change _(t-1)	1.154**	0.873*	0.690	0.989*

⁷ *The French Front National at 40: an Extremis Project eBook*. Edited by Dominic Ashton and Sarah Pilchick. 2013; <https://www.vox.com/2016/5/31/11722994/european-far-right-cas-mudde>

	(0.506)	(0.506)	(0.482)	(0.577)
Time	0.151	0.094	0.306*	0.038
	(0.177)	(0.173)	(0.164)	(0.212)
Female Leader _(t-1)	1.153	-0.188	-4.579	6.013*
	(3.504)	(3.428)	(4.034)	(3.334)
Cabinet Party _(t-1)	-3.511	-3.402	-2.617	-4.590
	(2.453)	(2.372)	(2.202)	(2.845)
Women in Parliament _(t-1)	0.461**	0.523***	0.230	0.577***
	(0.183)	(0.176)	(0.155)	(0.174)
Dis. Mag.	-0.007	-0.011	-0.004	-0.005
	(0.037)	(0.035)	(0.036)	(0.033)
PR	1.093	-5.165	-4.184	-0.837
	(4.287)	(4.595)	(4.240)	(4.982)
Quota Law	6.561*	4.305	7.693*	7.148*
	(3.444)	(3.463)	(4.123)	(3.710)
Western Europe	-2.215	-3.782		-5.736
	(3.908)	(3.741)		(3.345)
M/F Ratio _(t-1) *Vote Change _(t-1)	-0.868***	-0.688**	-0.569*	-0.798**
	(0.317)	(0.318)	(0.302)	(0.365)
Constant	-294.866	-179.006	-600.021*	-67.074
	(353.765)	(344.884)	(327.862)	(423.137)
Random-effect party	0	0	0	0
Random-effect country	23.91	20.13	26.03	0
Random-effect residual	39.01	36.98	29.48	66.21
N	58	58	58	58
Log Likelihood	-197.883	-195.299	-187.461	
AIC	425.765	422.599	414.922	
BIC	456.672	455.566	456.131	
Fixed effect for France	No	Yes	Yes	No
Fixed effects for Croatia, Greece, Italy, Austria, Denmark (the remaining 5 influential countries for full Model 5)	No	No	Yes	No

Note: Results are based on multilevel analyses with random intercepts for the country and party levels of the data. The dependent variable is the percentage of women among the RRP parties' MPs in national, lower-chamber legislature. Standard errors in parentheses. Models 2 and 3 using fixed effects for influential countries and adapted intercept were estimated using the influence.ME package for R (Nieuwenhuis et al. 2012). Model 4, robust regression, was estimated using robustlmm package for R (Koller 2016).

***p < .01; **p < .05; *p < .1

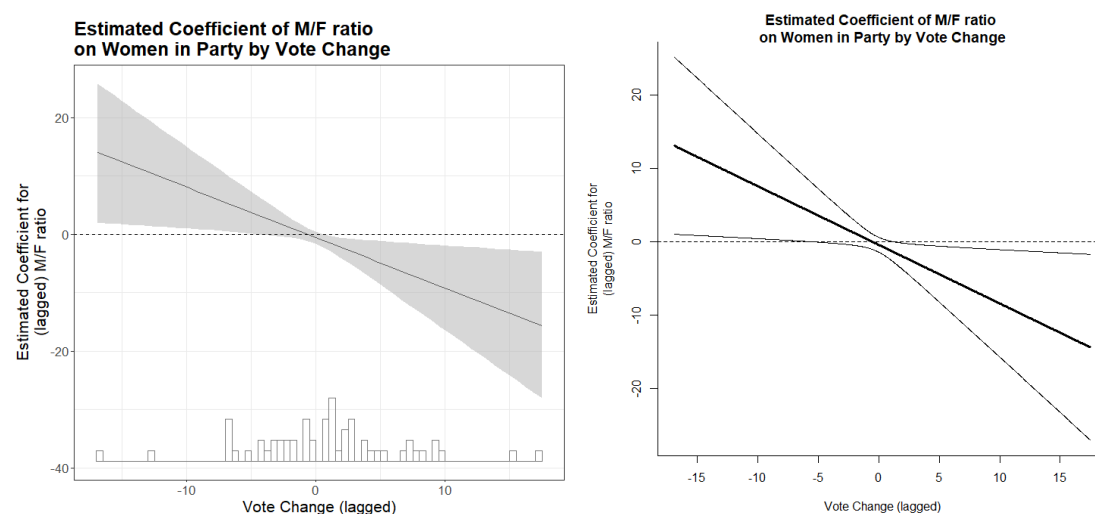
Model 1 of Table 1 shows our original specification with all controls (Table 1 Model 5, Weeks et al. 2023 p. 9). Models 2 and 3 of Table 1 show the preferred handling technique of Van der Meer et al. (2010), which is to include fixed-effect dummy variables at higher levels and set the intercept vector to value 0 for observations within the influential higher level units (see also Langford and Lewis 1998:125). Model 2 excludes the influence of France, speaking to the particular concern about the influence of FN 2012. It shows an interaction coefficient of -0.69, significant at the 0.05 level. However, as Ruiter and de Graf (2010) assert, this stopping criterion is arbitrary. A more rigorous approach to test the robustness of findings is to apply handling techniques to all cases found to be influential. Model 3 therefore excludes the influence of all 6 countries identified as influential. Model 3 shows an interaction coefficient of -0.57, significant at the 0.1 level ($p = 0.0649$).

Another approach to handling influential cases is robust regression. Robust regression assigns robustness weights to different observations, and recently approaches have been developed to apply robust regression to multilevel models (Koller 2016). Model 4 of Table 1 uses robust regression while still respecting the hierarchical nature of our data. It assigns robustness weights for 10 residuals, ranging from 0.436 to 0.988; the remaining weights are equal to 1. The case of FN 2012 is assigned a weight of 0.687. The output of Model 4 is again very close to our original specification, showing an interaction coefficient of -0.798, standard error of 0.365, and significant at conventional levels.⁸

In the right panel of Figure 2, we present the marginal effects plot from the newly estimated robust regression model (Model 4). Compare this to the original Figure 5 from Weeks et al. (2023) on the left. As seen in both, radical right populist parties respond with strategic increases in women's descriptive representation under conditions of electoral threat and a men-dominated electorate.

⁸ In addition, we estimate our original specification with all controls (Model 1 of Table 1 in this note) using least absolute deviation, another technique recommended to handle outliers (e.g., Aguinis et al. 2013). Unlike OLS, least absolute deviation chooses values of the regression coefficients that minimize the sum of the absolute values of the residuals. Our results are robust to this specification. The model returns an interaction coefficient of -0.401 with lower bound -1.419 and upper bound -0.258 (estimated using the quantreg package in R, Koenker et al. 2018). However, because least absolute deviation does not account for the nested structure of our data, we prefer the other two approaches described.

Figure 2: Marginal Effects Plot, Comparing Weeks et al. 2023 Figure 5 to Robust Regression (Table 1 Model 4)



The fixed effects and robust regression models estimated in Table 1 above (Models 2 – 4) and in Appendix Table B1 use restricted (or residual) maximum likelihood (REML) estimation, rather than MLE as we use in our main analysis in the original article. There is no option to use MLE for the handling techniques shown here, which are estimated in the influence.ME and robustlmm packages. However, to ease any concerns that the comparability of our results shown above could be sensitive to MLE versus REML estimation methods, we re-estimate models from our main analysis (Tables 1 and 2 of Weeks et al. 2023) using REML, the preferred estimation of GJ. As Table B1 of the Appendix shows, our findings are robust to REML estimation, and thus, we remain confident in their comparability to the results of the handling techniques presented in Table 1. We discuss this modeling decision further in the subsequent section on measurement and modeling strategies.

In sum, employing a variety of different estimation procedures for models with influential observations, we obtain coefficient estimates that are statistically and, importantly, substantively similar to our original findings that RRP parties see higher percentages of women MPs when they are both electorally weak and have a significant gender gap of untapped female electorate.

The Use of Strategic Descriptive Representation by the Front National in 2012

To quote statistician Stephen Samuels, “For covariates, one wants to identify not outliers per se, but those with high leverage. But the decision about what to do with these

(influential observations) is not automatic; sometimes they are the most important points and must be kept.”⁹

The FN case is one such observation. As discussed above, as one of the oldest and most electorally successful RRP parties and a party led by a woman, the FN has significant importance for the study of RRP parties in general and for the study of the use of gender as an electoral strategy in particular. It is therefore important to keep it in the analysis and more carefully examine it. In this next section of our response, we present the case of the Front National in 2012 as a third illustrative case study of strategic descriptive representation in action. Our original article combines cross-national analysis with two typical, “on-the-line”, cases for qualitative study: PVV in 2017 and SVP in 2015. The case of Front National in 2012 is not in our table of robust typical cases (see Table A7 in the Supplementary Information) because the residuals fall outside of the one standard deviation cutoff.¹⁰ In this case, the actual increase in the share of women in office was larger than the predicted increase from our models. Instead, as shown by the Cook’s d statistics, it can be considered an influential case, one that has the potential to contribute strongly to the probability of the hypothesis being true.

Recent qualitative research makes the argument that statistically influential cases like the FN in 2012 make excellent case studies for the goal of identifying causal effects. Herron and Quinn (2016) evaluate nine case selection methods discussed at length by Seawright and Gerring (2008). Their simulation-based results show that influential cases have the strongest performance on a number of criteria for case selection for the goal of causal inference.

In contrast to GJ’s assertion, a close look at the electoral context facing the FN in the run up to the 2012 elections suggests that not only is this observation statistically influential, it is also *substantively important for understanding and illustrating the theory of strategic descriptive representation*. First, the FN entered the run up to the 2012 parliamentary elections under the perfect conditions for strategic descriptive representation. They had lost almost 7 percentage points in the previous 2007 election. Their electorate was overwhelmingly made up of men, with a male-female voter ratio in 2007 of 3.92, almost twice as much as the mean ratio across the sample. Given these electoral conditions, our theory would strongly predict that they adopt a strategy of boosting women MPs. A closer examination of the case reveals just this strategic approach on the part of the party.

⁹ “Outliers,” <https://www3.nd.edu/~rwilliam/stats2/l24.pdf>, April 2016: page 15.

¹⁰ The residual values fall between 8.33 and 10.27 depending on the model, slightly above our one standard deviation cutoff of 8.3. The FN 2012 case would, however, be considered a “typical” case according to Lieberman’s (2005: 445) wider criteria of those cases with residuals falling within two standard deviations of the mean. Rohlfing and Starke (2013: 503) similarly find the one standard deviation cutoff to be too stringent.

French electoral politics is special within the set of comparative European cases based on its national adoption of gender parity laws on top of highly disproportional electoral rules.¹¹ With the percentage of women parliamentary candidates largely fixed in recent times at close to 50% and the likelihood of electoral success for this niche party fairly low in any plurality-based election up until this time, it becomes important to consider the FN's placement of women candidates within potentially winnable districts to assess to what extent strategic descriptive representation was pursued. Consistent with our predictions, in 2012, the FN prioritized the placement of women in electorally advantageous districts, to an extent not seen in previous elections. In 2012, women candidates were run in six of the top ten districts based on the districts' previous FN legislative electoral performance.¹² The number is five out of the top ten based on the districts' FN performance in the previous (2012) presidential election. In contrast, in 2002, when the Party was electorally strong and had a much lower male-female voter ratio (1.4) - and our theory would *not* expect them to pursue inclusive strategic descriptive representation - only three of the top ten best seats were contested by women.¹³

This promotion of women to the most winnable districts is particularly significant for a party historically likely to only win a few seats at most,¹⁴ and it reflected a deliberate strategy of emphasizing women to “soften” the FN image in order to attract women voters (Mondon 2014). For example, in tweets, the FN emphasized how mainstream parties were running fewer women MP candidates across the country than the FN, this despite the implementation of stronger national parity law provisions for the 2012 elections.¹⁵ This publicity strategy was complemented by a visual strategy of promoting women over men in the FN party campaign literature. As Dumitrescu (2017) finds, in the 2012 campaign materials, women candidates were presented using larger photos that were also zoomed in closer on the women's faces than those of the FN men candidates. Designed to increase the perceived approachability of candidates (Dumitrescu 2017; Kress and Van Leeuwen 2006) and thus, to use women MP candidates to attract voter support, this gendered visual tactic represented a shift in the FN's publicity materials over previous elections.

¹¹ Both of these factors are controlled for in the regression models.

¹² Presidential and legislative election data by constituency, CDSP: Ressources en ligne (sciences-po.fr).

¹³ This number holds whether one measures top seats based on previous legislative or presidential results.

¹⁴ Prior to the 2012 elections, the mean seat attainment under plurality rules for the FN in legislative elections was 0.25 seats; since 1973, there were six elections in which they won no seats and two in which they won 1.

¹⁵ Despite their prior opposition to the parity law, the FN, under the gender-prioritizing campaign of Party Leader Marine Le Pen, became one of the best performers in running women candidates in elections (Childs and Murray 2014). May 21, 2012 tweets referencing Jean-Baptiste Garat, “Plus de 70% des candidats UMP sont des hommes.” *Le Figaro*, May 21, 2012.

In addition, the FN placed a woman - Marion Maréchal-Le Pen (granddaughter of FN founder Jean-Marie Le Pen and niece of Marine Le Pen) - in their **most** electable district based on prior 2012 FN presidential support. Maréchal-Le Pen had not contested that seat before, and indeed, she *replaced* a man FN candidate (Guy Macary) who had been popular and long contested that seat. In other words, this was not a haphazard or random behavior with minimal costs; it was a strategic choice to reduce the number of men in an electable district and boost the number of women. The national news media picked up on this emphasis of women by the FN, not only highlighting that the party was putting a woman in their most favored district, but playing up her identity as a woman in articles. Being a young, attractive woman, Marion Maréchal-Le Pen fit within the FN campaign strategy of de-demonization and modernization, and the rebranding process of the FN following Marine Le Pen's leadership.¹⁶ Marion Maréchal-Le Pen, herself, also highlighted her gender, by saying how being a woman - as a symbol of equality - could help her win the seat.¹⁷ By presenting a young woman, the FN also hoped to attract young women voters, who were underrepresented among the FN's electorate (Mayer 2013a).

And the FN's use of strategic descriptive representation - of introducing more women to winnable districts and to more visible positions - seems to have been successful. The FN succeeded in getting two MPs elected, one of them, a woman - Marion Maréchal-Le Pen. Women voters, in particular young women, increased their support of the FN to the point that scholars such as Nonna Mayer (2013a; 2013b; 2015) dubbed 2012 a "turning point" election year in which the gender gap in support for the FN ceased to exist. The FN's vote share subsequently increased by 9.3%. Consistent with our theory and the data we present above, Dumitrescu (2017) attributes this attraction of more voters, and more women in particular, to the FN's deliberate campaign strategy prioritizing women MPs. In summary, our examination of the 2012 French FN case suggests that it is not a mere statistical nuisance, but instead exemplifies the phenomenon of strategic descriptive representation that our article seeks to understand.

In their rejoinder, GJ challenge our discussion of the substantive importance of the French FN case. They offer additional statistical analyses focusing on the placement of women among *all* FN parliamentary candidates. However, our claim above is not that all women candidates are elevated to better districts under strategic descriptive representation strategies. We make the case, instead, that it is necessary to consider the

¹⁶ Although, it is important to note that many scholars (Crépon 2012, 2015; Dézé 2012; Fourest and Venner 2012; Dumitrescu 2017) have found few substantive policy changes and shifts in the 2012 FN manifesto. The 2012 strategy is that of descriptive representation of women over substantive representation.

¹⁷ Gérard Bon and John Irish, "Le Pen's granddaughter to run for French parliament." France Election News. April 25, 2012
<https://www.reuters.com/article/us-france-election-lepen-idUSBRE83O0VV20120425>

placement of women among the set of *feasible* parliamentary candidates. This is especially true in the context of the French FN case, a men-dominated party which rarely gains seats in a country with a candidate gender parity law. That is why our discussion of the FN case above focuses on the small set of feasible districts for seat attainment. The vast majority of districts were uncompetitive for the FN in 2012 and previous years. Further, GJ do not engage with the case-based literature we discuss, which offers rich detail about the gendered nature of the party's campaign strategies in this election. For example, studies show that the FN's 2012 campaign was characterized by a new emphasis on women in visual depictions and media (Dumitrescu 2017; Mayer 2013a), in line with our theory. Together our discussion shows that the FN 2012 case exemplifies strategic descriptive representation: replacing a long-term man candidate with a well-known woman candidate in one of the few competitive FN seats, amid national-level vote loss and a significant untapped women voter base.

Measurement and modeling strategies

GJ make three main critiques related to our measurement approach and modeling strategies. We address each in turn, below.

Multilevel Modeling Decisions

GJ question our choice to model women's descriptive representation using multilevel models with maximum likelihood rather than restricted maximum likelihood (REML). REML is the default for STATA and, as a result, it is the approach used in many multilevel models employing STATA code published in political science journals. The reason we chose maximum likelihood is simple. We began the analysis in STATA using their default settings and only later transitioned to R. More importantly, as GJ show in Figure 2 and we show in Appendix Table B1, the choice about using REML is inconsequential to our findings. As both we and GJ report, in our fully specified model which includes country- and party-level controls (M5), the interaction coefficient remains statistically significant at the 0.05 level.

GJ note that the use of maximum likelihood via the lme4 package in R returns a warning message about boundary fit which could suggest the model fails to converge. To confirm that this issue does not impact our results, we re-run all specifications in Table 1 of the article's main text using: 1) linear panel model estimation, and 2) Bayesian multilevel regression model estimation. First, we use the plm package in R to fit linear panel models with random effects (see Table B2 of Appendix). These models employ Swamy-Arora's transformation, a preferred estimator for small samples (Baltagi and Chang 1994), and the boundary fit issue is not relevant to these specifications.

Reassuringly, our main results hold. The only exception is that in Model 3 the p-value for the key interaction term (male/female voter ratio and vote change lagged) is 0.05 (compared to $p=0.04$ for the same term in our published text). Second, we re-run all specifications in Table 1 using the `rstanarm` package for R, which employs a Bayesian hierarchical regression framework. The Bayesian framework has the advantage of offering greater prediction accuracy compared to the same models fitted in a frequentist framework (Gelman and Hill 2006). Table B3 in the Appendix shows the median of the posterior distribution for each variable and its credible interval.¹⁸ Again, our results are largely robust, with the size and direction of coefficients quite similar to the other specifications that we run. 95% credible intervals rather than p-values are reported in Bayesian output, but we note that the key interaction term remains negative and statistically significant (i.e., credible in Bayesian terms) across all models with controls, and falls just short of conventional significance in Model 2 (no controls).¹⁹ In other words, our results are robust to different multilevel modeling assumptions.

Multilevel Models vs. Fixed Effects

In our article, we employ multilevel models with observations nested in parties and countries. As GJ acknowledge on page 13, this model choice is "theoretically justified." In the appendix of the article, we rerun the analyses with OLS models in order to test the robustness of the findings. This is not our preferred specification because it ignores the fact that the data consists of countries within which are nested multiple RRP parties contesting multiple elections. Multilevel random-intercept models account for the nested structure of our data and take into consideration both within-level and between-level variation in the dependent variable (Gelman and Hill 2007). This choice also allows us to include important party- and country-level variables that are central to existing theories of women's descriptive representation.

Country fixed effects control for all time-invariant differences among countries, estimating "within" effects. This comes at the cost of estimating "between" effects, differences across countries. Were we to include a fixed effects black box for country impact, we would not be able to comment on the impact of variables key to the descriptive representation literature such as electoral rules and the differential historical experiences of countries in Eastern and Western Europe. Many countries in our dataset have also never seen a woman leader of an RRP party or adopted a national quota law. We note that in their replication, GJ include covariates that vary across countries but not

¹⁸ We fit the model using 4 chains for 14,000 iterations, in which the first 7,000 are the burn-in period, and the last 7,000 are the sampling period. We run model diagnostics to check for convergence issues for all the models, and these showed no evidence of convergence issues.

¹⁹ When we use a 90% credible interval, the interaction term in Model 2 is significant.

over time (PR electoral system and Western Europe), which should be dropped in any model including country fixed effects.

We confirm the appropriateness of our empirical strategy by performing Hausman tests. For each specification in Table 1, our test fails to reject the null that the preferred model is random effects (versus the alternative, country fixed effects). Note that in these tests we drop PR electoral system and Western Europe variables from Model 5 due to the collinearity issue on inclusion of country FE. Given the results of the Hausman tests, and noting that we are speaking to a well-established literature which has identified important structural, time-invariant variables as key determinants of women's representation, our modeling choice is well-justified.

Measurement of Explanatory Variable: M/F Voter Ratio vs. Share Men Voters

Lastly, GJ advance a new operationalization of our main explanatory variable of the gender gap of men and women voters. They claim that the use of a ratio to measure gender gaps is not “intuitive” and instead use our male/female voter ratio to calculate the share of men voters. However, they misunderstand our data, and as a result their new variable does not measure what they claim. Our male/female ratio variable is calculated as the percentage of men who voted for the RRP party among all men voting for a party in that country's election over the percentage of women in the country voting for the RRP party among all women voting for a party in that country's election. It thus captures whether men are disproportionately voting for a particular party relative to women or, in other words, the relative scale of difference in support. It is not, as GJ assert, a measure of what percentage of a *given RRP party's vote share* is composed of men (versus women) voters; that would require different data. Therefore, their analysis does not provide an alternative test of our theory about the impact of untapped women voters.

Moreover, the operationalization of the gender gap in voting we employed follows a long line of studies in gender and politics which measure gender gaps as ratios, including for example male/female labor force participation, male/female educational achievement, male/female candidate ratios, and male/female political speech ratios (e.g., Matland 1998; Stockemer 2008; Giger et al. 2014; Wahman et al. 2021), although we are among the first to do this for party voting, in part because such data has yet to be systematically compiled (but see Donovan 2023). The annual Global Gender Gap Index (World Economic Forum) also converts all data into ratios, “in order to ensure that the Index is capturing gaps between women and men's attainment levels, rather than the levels themselves” (Global Gender Gap Report 2023, p. 64). It is thus a conventional measurement approach that fits our theoretical aims well.

Conclusion

Taken together, the statistical and case study results support the reliability of our original findings. As GJ report, computer reproducibility is nearly perfect, and our main results are verified bar one star sign on one coefficient, due to an anomaly in the Stargazer package for R. This issue does not affect the vast majority of our tables and thus does not change our findings in any substantive way. When influential case diagnostics are systematic, and appropriate models recognizing the multilevel structure of our data are employed, the results of our models are robust to the removal of influential cases and, more importantly, to the recommended robust regression handling techniques. Further examination of the 2012 French FN case suggests that this is an important case that should not be ignored, but instead sheds further light on the phenomenon of strategic descriptive representation that our article seeks to understand. With regard to questions of measurement and modeling technique, we follow conventional approaches, driven by our theory and the structure of our data. Our use of multilevel models is justified given the nested structure of our data and the theoretically important country-level covariates we include, and Hausman tests confirm that it is the more appropriate choice (compared to country fixed effects, which GJ suggest). The use of REML vs. MLE makes no difference to our findings, and our results are also robust to both linear and Bayesian regression estimations for multilevel models. And new efforts by GJ to present a different operationalization of our main explanatory variable misunderstand our data and thus do not measure the party-level share of men voters they aim to capture. In conclusion, the evidence we present here reconfirms our article's main theory: electorally weak RRP parties increase the percentage of women MPs when they lack women voters.

References

- Aguinis, H., Gottfredson, R.K. and Joo, H., 2013. Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, 16(2), pp.270-301.
- Baltagi, B. H., and Y.-J. Chang. 1994. Incomplete panels: A comparative study of alternative estimators for the *unbalanced one-way error component regression model*. *Journal of Econometrics*, 62, pp.67–89.
- Belsley, D. A., Kuh, E. and Welsch, R.E., 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley.
- Bouvet, L., 2013. “Foreword,” in Ashton, D. and Pilchick, S (eds) *The French Front National at 40*. Extremis Project. Ebook. p. 3.
- Childs, S., and Murray, R., 2014. ‘Feminising political parties’ in Childs, S.L. and Campbell, R. (eds) *Deeds and words: Gendering politics after Joni Lovenduski*. ECPR Studies in European Political Science. pp. 73-90.
- Crépon, S., 2012. *Enquête au coeur de nouveau Front national*. Paris: Nouveau Monde éditions.
- Crépon, S., 2015. La politique des moeurs au Front national. In Crépon, S., Dézé, A. and Mayer, N. (eds) *Les faux-semblants du Front national: Sociologie d’un parti politique*. Paris: Presses de Sciences Po. pp. 185-206.
- Dézé, A., 2012. *Le Front national: à la conquête du pouvoir?* Paris: Armand Colin.
- Donovan, T., 2023. Measuring and predicting the radical-right gender gap, *West European Politics*, 46(1), pp. 255-264.
- Dumitrescu, D., 2017. Up, close and personal: The new Front National visual strategy under Marine Le Pen. *French Politics*, 15(1), pp. 1-26.
- Fourest, C. and Venner, F., 2012. *Marine Le Pen démasquée*. Paris: Le Livre de Poche.
- Gelman, Andrew and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel Hierarchical Models*. Cambridge University Press.
- Gerring, J. and Cojocaru, L., 2016. Selecting cases for intensive analysis: A diversity of goals and methods. *Sociological Methods & Research*, 45(3), pp.392-423.
- Giger, N., Holli, A.M., Lefkofridi, Z. and Wass, H. 2014. The gender gap in same-gender voting: The role of context. *Electoral Studies*, 35, pp.303-314.

The Global Gender Gap Report 2023. World Economic Forum. Available on: <http://reports.weforum.org/globalgender-gap-report-2023> [retrieved on 14 February 2024]

Good, I.J., 1982. Standardized tail-area probabilities. *Journal of Statistical Computation and Simulation*, 16(1), pp. 65-66.

Herron, M.C. and Quinn, K.M., 2016. A careful look at modern case selection methods. *Sociological Methods & Research*, 45(3), pp.458-492.

Koller, M., 2016. robustlmm: an R package for robust estimation of linear mixed-effects models. *Journal of statistical software*, 75, pp.1-24.

Koenker, R., Portnoy, S., Ng, P.T., Zeileis, A., Grosjean, P. and Ripley, B.D., 2018. Package ‘quantreg’. Available at: R-CRAN: <https://cran.rproject.org/web/packages/quantreg/quantreg.pdf>

Kress, G. and Van Leeuwen, T., 2006. *Reading Images: The Grammar of Visual Design*. New York: Routledge.

Langford, I. H. and Lewis, T., 1998. Outliers in Multilevel Data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 161, pp.121–60.

Lieberman, E., 2005. Nested Analysis as a Mixed-Member Strategy for Comparative Research. *American Political Science Review*, 99(3), pp.435-52.

Matland, R.E. 1998. Women's representation in national legislatures: Developed and developing countries. *Legislative Studies Quarterly*, 23(1), pp.109-125.

Mayer, N., 2013a. ‘The Far Right in France and the Gender Gap’ in Ashton, D. and Pilchick, S. (eds) *The French Front National at 40*. Extremis Project. Ebook, pp. 17-21

Mayer, N., 2013b. From Jean- Marie to Marine Le Pen: Electoral Change on the Far Right. *Parliamentary Affairs*, 66, pp. 160– 178.

Mayer, N., 2015. The Closing of the Radical Right Gender Gap in France? *French Politics*, 13(4), pp. 391– 414.

Mondon, A., 2014. The Front National in the Twenty-First Century. *Modern and Contemporary France*, 22(3), pp. 301-320.

Murray, R., 2013. Toward Parity Democracy? Gender in the 2012 French Legislative Election. *Parliamentary Affairs*, 66, pp. 197-212.

- Nieuwenhuis, R., Te Grotenhuis, H.F. and Pelzer, B.J., 2012. Influence. ME: tools for detecting influential data in mixed effects models. *The R Journal*, 4(2), pp.38-47.
- Rohlfing, I., and Starke, P., 2013. Building on Solid Ground: Robust Case Selection in Multi-Method Research. *Swiss Political Science Review*, 19(4), pp.492-512.
- Ruiter, S. and De Graaf, N.D., 2010. National Religious Context and Volunteering: More Rigorous Tests Supporting the Association. *American Sociological Review*, 75(1), pp.179–84.
- Seawright, J. and Gerring, J., 2008. Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political research quarterly*, 61(2), pp.294-308.
- Stockemer, D. 2008. Women's representation: A comparison between Europe and the Americas. *Politics*, 28(2), pp.65-73.
- Van der Meer, T., Te Grotenhuis, M. and Pelzer, B., 2010. Influential cases in multilevel modeling: A methodological comment. *American Sociological Review*, 75(1), pp.173-178.
- Wahman, M., Frantzeskakis, N. and Yildirim, T.M.. 2021. From thin to thick representation: how a female president shapes female parliamentary behavior. *American Political Science Review*, 115(2), pp.360-378.
- Weeks, A.C., Meguid, B.M., Kittilson, M. and Coffé, H., 2023. When do Männerparteien elect women? Radical right populist parties and strategic descriptive representation. *American Political Science Review*, 117(2), pp.421-438.

Appendix	page
Appendix A: Influential Case Identification and Handling	2
Table A1: Determinants of Women's Representation in Radical Right Populist Parties, Removing Influential Country Cases.....	2
Figure A1: Diagnostics for Influential Cases in Multilevel Models.....	4
Table A2: Determinants of Women's Representation in Radical Right Parties, Removing Influential Country Cases (dropping PR electoral system).....	5
Appendix B: Additional Estimation Strategies	7
Table B1: Determinants of Women's Representation in Radical Right Populist Parties, REML...7	
Table B2: Determinants of Women's Representation in Radical Right Populist Parties, Linear Panel Models (plm).....	9
Table B3: Determinants of Women's Representation in Radical Right Populist Parties, Bayesian Estimations.....	10

Appendix A: Influential Case Identification and Handling

Below we replicate our results using the deletion process advocated by GJ, while additionally accounting for the multilevel data structure and excluding all influential cases per our argument in the body of our reply. We calculate the influential cases for Model 5, the fullest model which mitigates the potential for omitted variable bias (see Figure 1 in the text). Because excluding all influential cases reduces our observations significantly ($N=31$ versus the original $N=58$), we are not able to specify the full Model 5 with all controls while deleting these cases.

In light of this, we pursue two approaches to test the robustness of our argument via deletion.

In Table A1, we drop the six influential cases from our full Model 5 diagnostics and specify a model with no covariates. Model 1 of Table A1 presents the original results in our paper with all observations included (original Table 1 Model 2, p. 9), and Model 2 of Table A1 presents the results when dropping all influential cases detected. As the table shows, the results do not change: the interaction between vote change and M/F voter ratio is negative as expected. The coefficient is very similar in size compared with our original findings (-0.69 compared with -0.66 in our original results), and the standard error of 0.35 is also nearly the same as the 0.34 reported with all cases included.²⁰

Table A1: Determinants of Women's Representation in Radical Right Populist Parties, Removing Influential Country Cases

	Model 1 (original Table 1, Model 2)	Model 2 (influential cases removed)
M/F Ratio _(t-1)	-0.551 (0.478)	0.987 (0.979)
Vote Change _(t-1)	0.840 (0.538)	0.689 (0.504)
M/F Ratio _(t-1) *Vote Change _(t-1)	-0.663* (0.338)	-0.689* (0.353)
Constant	20.684*** (2.271)	18.150*** (2.848)
Random-effect party	1.14	45.20
Random-effect country	57.15	17.96
Random-effect residual	48.00	16.09
N	58	31
Log Likelihood	-208.362	-101.632

²⁰ If we take into consideration the much reduced sample size of the data included in Model 2 ($N=31$) and apply Good's (1982) adjustment for small sample sizes less than 100, the estimated p-value of the interaction term becomes 0.036. This is comparable to the Good's adjusted p-value for the interaction term in the full-sample ($N=58$) Model 1 above ($p=0.042$).

AIC	430.723	217.263
BIC	445.146	227.301

Note:

Results are based on multilevel analyses with random intercepts for the country and party levels of the data. The dependent variable is the percentage of women among the RRP parties' MPs in national, lower-chamber legislature. Standard errors in parentheses. Model 2 removing influential country cases was estimated using the influence.ME package for R (Nieuwenhuis et al. 2012).

***p < .01; **p < .05; *p < .1

As a second approach, we calculate the influential cases for and re-run the analysis on the exact same model specification. This is the approach advocated by GJ in their rejoinder, but we note that they do not follow their recommendations in their treatment of Model 5, perhaps for the same reason we do not in Table A1 above; it is not possible to run the originally-specified model minus influential cases on such a limited number of observations.

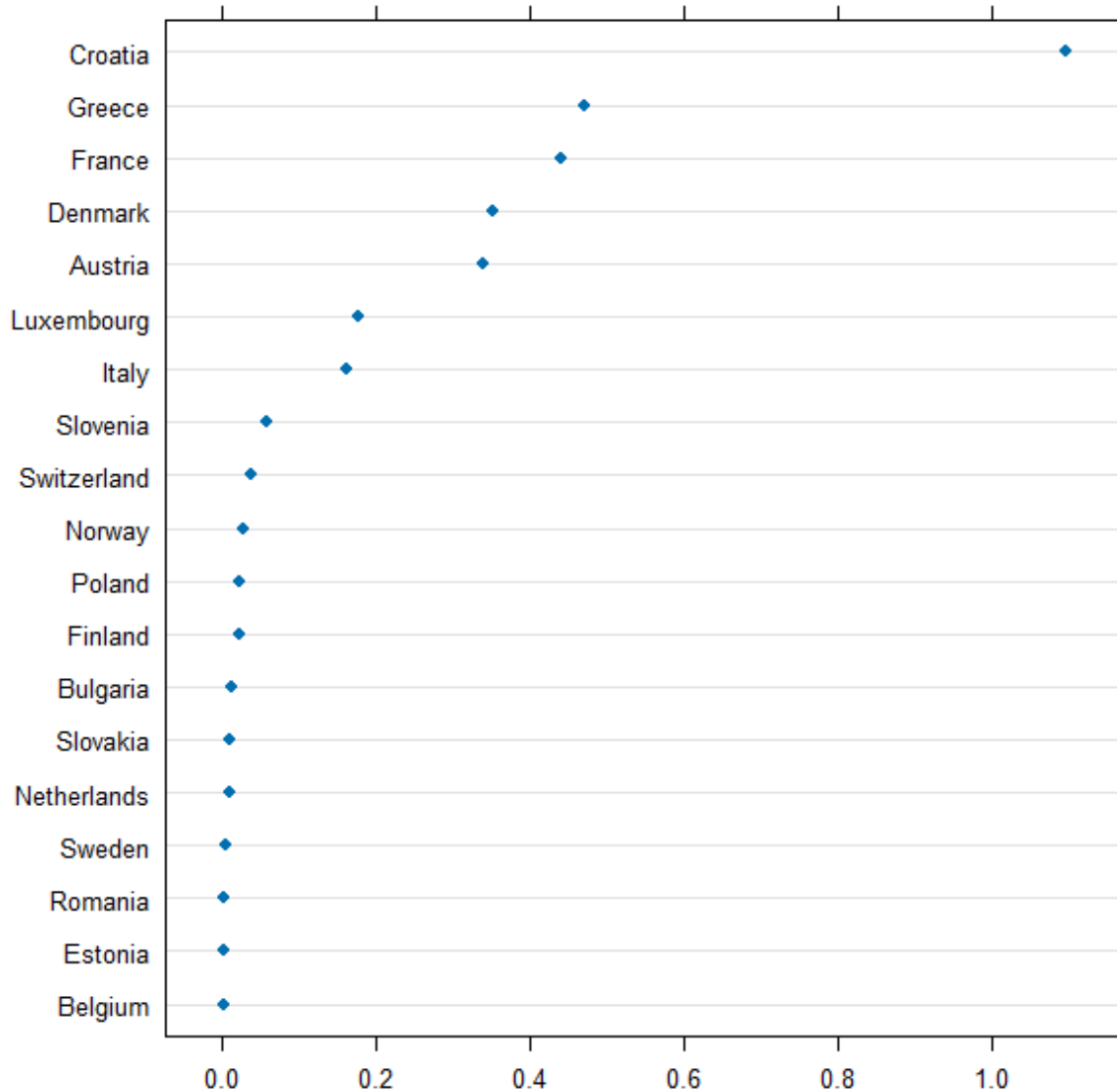
Starting again with Model 5, we identify that the variable PR electoral system limits our ability to run the model without the influential cases, because we are left without enough variation in the electoral system variable in the reduced data. Therefore, we re-estimate the Cook's D for a modified Model 5 which drops the PR electoral system variable. As shown in Figure A1, we find that five countries are influential for the modified model: Croatia, France, Greece, Austria, and Denmark. We then re-estimate the same model specification, dropping these cases.²¹ Model 1 of Table A2 presents the results of modified Model 5 with all observations included, and Model 2 of Table A2 presents the results when dropping all influential cases that pertain to this specification. As the Table shows, despite dropping 18 observations out of 58, the results are substantially the same: the interaction between vote change and M/F voter ratio is negative as expected (-0.65), significant at the 0.1 level.²²

While this is reassuring, we reiterate that best practices note that this is not an appropriate way to deal with (correctly-coded) influential observations. We show that our results are robust to various handling techniques for downweighting influential observations in Table 1 of the text.

²¹ Note that GJ drop the original six countries in their Model 5 in the rejoinder, not the five countries as their argument would require.

²² If we take into consideration the reduced sample size of the data included in Model 2, Table A2 (N=40) and apply Good's (1982) adjustment for small sample sizes less than 100, the estimated p-value of the interaction term becomes 0.049.

Figure A1: Diagnostics for Influential Cases in Multilevel Models: Cook's Distance for Modified Model 5



Note: Cook's distance values were calculated for Model 5 excluding the PR electoral system variable (see Table A1). Using the standard cutoff criterion of $4/n$ (where n = number of cases; Belsley et al. 1980), the cutoff for influence is $4/19 = 0.2105$. According to this cutoff, five country cases are influential: Croatia, Greece, France, Denmark, and Austria.

Table A2: Determinants of Women's Representation in Radical Right Parties, Removing Influential Country Cases (dropping PR electoral system)

	Model 1 (original Table 1, Model 5, dropping PR)	Model 2 (dropping PR and influential cases removed)
M/F Ratio _(t-1)	-0.582 (0.431)	1.213 (1.054)
Vote Change _(t-1)	1.125** (0.478)	0.503 (0.491)
Time	0.146 (0.177)	0.042 (0.192)
Female Leader _(t-1)	1.145 (3.426)	-2.577 (4.346)
Cabinet Party _(t-1)	-3.409 (2.401)	1.364 (2.137)
Women in Parliament _(t-1)	0.483*** (0.174)	0.748*** (0.154)
Dis. Mag.	-0.006 (0.036)	-0.026 (0.027)
Quota Law	6.489* (3.421)	6.689* (3.389)
Western Europe	-2.592 (3.746)	-8.881** (3.018)
M/F Ratio _(t-1) *Vote Change _(t-1)	-0.853*** (0.304)	-0.646* (0.356)
Constant	-284.994 (354.603)	-81.309 (383.965)
Random-effect party	0.00	0.00
Random-effect country	22.51	8.657
Random-effect residual	39.63	21.211
N	58	40
Log Likelihood	-197.911	-122.864
AIC	423.822	273.728

BIC	452.668	297.372
-----	---------	---------

Note:

Results are based on multilevel analyses with random intercepts for the country and party levels of the data. The dependent variable is the percentage of women among the RRP parties' MPs in national, lower-chamber legislature. Standard errors in parentheses. Model 1 is a modified version of Model 5 from the article, with the PR electoral system variable removed to allow it to run. Model 2 removing influential country cases was estimated using the influence.ME package for R (Nieuwenhuis et al. 2012).

*** $p < .01$; ** $p < .05$; * $p < .1$

Appendix B

In Table B1, we re-estimate Table 1 from our main analysis (Weeks et al. 2023) using REML rather than MLE. The table shows results consistent with our main analysis.

Table B1: Determinants of Women's Representation in Radical Right Populist Parties, REML

	Model 1	Model 2	Model 3	Model 4	Model 5
M/F Ratio _(t-1)	-0.555 (0.499)	-0.549 (0.497)	-0.629 (0.453)	-0.724 (0.461)	-0.672 (0.474)
Vote Change _(t-1)	-0.167 (0.182)	0.842 (0.557)	0.897* (0.516)	0.995* (0.543)	1.099* (0.564)
Time			0.479*** (0.156)	0.479*** (0.159)	0.187 (0.196)
Female Leader _(t-1)				0.067 (4.033)	0.323 (4.020)
Cabinet Party _(t-1)				-4.006 (2.619)	-3.398 (2.710)
Women in Parliament _(t-1)					0.401* (0.213)
Dis. Mag.					-0.007 (0.045)
PR					1.850 (4.753)
Quota Law					6.563 (3.941)
Western Europe					-1.216 (4.721)
M/F Ratio _(t-1) * Vote Change _(t-1)		-0.664* (0.350)	-0.669** (0.323)	-0.764** (0.341)	-0.831** (0.352)
Constant	21.333*** (2.438)	20.680*** (2.363)	-941.860*** (312.709)	-941.144*** (320.223)	-367.172 (391.613)
Random-effect party	0	1.97	0	0	0
Random-effect country	71.06	61.48	60.83	55.22	41.91
Random-effect residual	52.71	51.23	43.15	44.24	45.27
N	58	58	58	58	58
Log Likelihood	-209.082	-207.483	-204.034	-198.697	-190.857
AIC	430.164	428.966	424.068	417.394	411.713

BIC	442.527	443.389	440.552	437.999	442.620
-----	---------	---------	---------	---------	---------

Note: Results are based on multilevel analyses with random intercepts for the country and party levels of the data. REML is employed. The dependent variable is the percentage of women among the RRP parties' MPs in national, lower-chamber legislature. Standard errors in parentheses.

*** $p < .01$; ** $p < .05$; * $p < .1$

Table B2: Determinants of Women's Representation in Radical Right Populist Parties, Linear Panel Models (plm)

	Model 1	Model 2	Model 3	Model 4	Model 5
M/F Ratio _(t-1)	-0.185 (0.601)	-0.344 (0.592)	-0.529 (0.580)	-0.556 (0.548)	-0.517 (0.572)
Vote Change _(t-1)	-0.136 (0.185)	0.883 (0.570)	0.877 (0.535)	1.172** (0.567)	1.170** (0.574)
Time			0.489*** (0.184)	0.454** (0.186)	0.135 (0.215)
Female Leader _(t-1)				3.948 (3.777)	2.211 (3.771)
Cabinet Party _(t-1)				-4.543* (2.730)	-3.569 (2.752)
Women in Parliament _(t-1)					0.468** (0.205)
Dis. Mag.					-0.009 (0.044)
PR					1.786 (4.872)
Quota Law					5.975 (4.023)
Western Europe					-1.931 (4.480)
M/F Ratio _(t-1) * Vote Change _(t-1)		-0.676* (0.359)	-0.647* (0.336)	-0.856** (0.358)	-0.874** (0.359)
Constant	20.534*** (2.488)	20.280*** (2.422)	-961.601*** (368.952)	-890.692** (373.247)	-263.045 (430.787)
R-squared	0.138	0.181	0.269	0.316	0.429
Adjusted R-squared	0.106	0.135	0.214	0.236	0.293
N	58	58	58	58	58

Note: Results are based on multilevel analyses with random intercepts for the country and party levels of the data. The dependent variable is the percentage of women among the RRP parties' MPs in national, lower-chamber legislature. Standard errors in parentheses.

***p < .01; **p < .05; *p < .1

The interaction term in Model 3 is significant at p=.054.

Table B3: Determinants of Women's Representation in Radical Right Populist Parties, Bayesian Estimations

	Model 1	Model 2	Model 3	Model 4	Model 5
<i>Fixed Effects</i>					
	<i>median/sd</i>	<i>median/sd</i>	<i>median/sd</i>	<i>median/sd</i>	<i>median/sd</i>
M/F Ratio _(t-1)	-0.43 (0.56)	-0.47 (0.55)	-0.59 (0.50)	-0.67 (0.51)	-0.59 (0.55)
Vote Change _(t-1)	-0.15 (0.18)	0.85 (0.57)	0.88 (0.52)	1.00 (0.57)	1.07 (0.58)
Time			0.49* (0.17)	0.48* (0.17)	0.17 (0.22)
Female Leader _(t-1)				0.87 (4.75)	1.01 (4.25)
Cabinet Party _(t-1)				-4.03 (2.71)	-3.42 (2.76)
Women in Parliament _(t-1)					0.43 (0.24)
Dis. Mag.					-0.01 (0.05)
PR					1.51 (5.11)
Quota Law					6.32 (4.01)
Western Europe					-1.74 (4.95)
M/F Ratio _(t-1) * Vote Change _(t-1)		-0.65 (0.36)	-0.65* (0.33)	-0.76* (0.36)	-0.82* (0.36)
Constant	20.99* (2.52)	20.50* (2.40)	-957.61* (343.74)	-950.90* (346.28)	-338.50 (444.13)
<i>Sigma</i>	<i>7.34*</i>	<i>7.23*</i>	<i>6.66*</i>	<i>6.81*</i>	<i>6.93*</i>

	(0.84)	(0.85)	(0.78)	(0.86)	(0.90)
<i>Random Effects</i>					
	<i>Sd</i>	<i>Sd</i>	<i>sd</i>	<i>sd</i>	<i>sd</i>
Country (Intercept)	7.52	7.12	7.32	6.77	5.76
Party (Intercept)	4.85	4.63	4.07	3.93	3.69
Residual	7.42	7.32	6.74	6.90	7.01
Num. levels: party 22, country 19					

*0 outside 95% credible interval (*ci* = *credible interval*; *sd* = *standard deviation*)