INSTITUTE for REPLICATION

No. 41
I4R DISCUSSION PAPER SERIES

On the Empirical Validity of "Gendered Reactions to Terrorist Attacks Can Cause Slumps not Bumps" (Holman et al., 2022)

Michael Jetter

Kieran Stockley

June 2023

I4R Discussion Paper Series

I4R DP No. 41

On the Empirical Validity of "Gendered Reactions to Terrorist Attacks Can Cause Slumps not Bumps" (Holman et al., 2022)

Michael Jetter[1], Kieran Stockley[2]

[1]University of Western Australia and Centre of Business Data Analytics (UWA), Crawley/Australia, IZA Bonn/Germany, CESifo Munich/Germany
[2]University of Western Australia, Crawley/Australia

June 2023

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

On the empirical validity of “Gendered reactions to terrorist attacks can cause slumps not bumps” (Holman et al., 2022)*

Michael Jetter† Kieran Stockley‡

May 13, 2023

Abstract

Holman et al. (2022; HMZ) propose women (compared to men) political leaders experience significant drops in public approval ratings after a transnational terrorist attack. After documenting how survey-based evaluations of then-Prime Minister Theresa May suffered after the 2017 Manchester Arena attack, HMZ assemble a country-quarter level panel database to explore the generality of their hypothesis. They report evidence suggesting women (compared to men) leaders systematically experience decreased public approval rates after major transnational terrorist attacks (p-value of 0.020). We find that result disappears once *any* of the following adjustments is implemented: (i) excluding election quarter covariates ($p = 0.104$); (ii) correcting objective coding errors in the election quarter covariates ($p = 0.058$); (iii) excluding the May-Manchester observation ($p = 0.098$); or (iv) clustering standard errors at the country level ($p = 0.558$). Exploring all 2^5 combinations of the five control groups HMZ incorporate in their specification, none of them clears the 5% threshold of statistical significance once the corrected election quarter variables are employed. We conclude that the empirical evidence does not provide sufficient support for HMZ’s abstract claim that “conventional theory on rally events requires revision: women leaders cannot count on rallies following major terrorist attacks.”

*We are thankful to Tushar Bharati and Leandro Magnusson for helpful feedback and discussions. All remaining errors are our own.

†University of Western Australia, Centre of Business Data Analytics (UWA), IZA (Bonn), and CESifo (Munich); 35 Stirling Highway, Crawley 6009, WA; email: mjetter7@gmail.com.

‡University of Western Australia; 35 Stirling Highway, Crawley 6009, WA; email: kieran.stockley@gmail.com.

1 Introduction

It is important to understand whether electorates systematically judge women leaders differently than men leaders. In this context, Holman et al. (2022; HMZ from hereon) present empirical evidence to suggest rally 'round the flag effects may not apply to women leaders after their polity experiences a major transnational terrorist attack. First, in a case study analysis, HMZ identify a statistically significant drop in then-Prime Minister Theresa May's approval ratings after the 2017 Manchester arena bombing. Second, HMZ assemble a panel database at the country-quarter level to explore whether women (compared to men) political leaders *generally* suffer from diminished approval ratings after their country experiences a major transnational terrorist attack. Presenting estimates from one regression in the main body of the paper, HMZ derive a negative correlation between women leadership and approval rates after such events that is statistically significant at the 5% level.

It is this second, global setting our replication focuses on. In the following pages, we first show that the May-Manchester results remain robust, even though we identify a coding error that incorrectly attributes 35,060 responses (equivalent to 36% of all observations) in survey Waves 7 and 8 to Theresa May, rather than David Cameron who was prime minister at the time. In the main part of the paper, we then explore HMZ's country-quarter level panel analysis. To preview our findings, we conclude HMZ's generalized claim that "women leaders cannot count on rallies following major terrorist attacks" is not empirically supported. We reach this conclusion through various alternative specifications that, independent of each other, yield statistically insignificant estimates at conventional levels.

HMZ's country-quarter level database connects the Executive Approval Dataset (Carlin et al., 2019) with the Global Terrorism Database (START, 2022; GTD), among other sources. This produces a database of 4,328 observations in 44 countries from 1975 to 2017. Presenting estimates from one regression that accounts for several sets of control variables and country-fixed effects, HMZ find the coefficient associated with the interaction term between featuring a woman leader and experiencing a major transnational terror attack in the preceding quarter to be negative and statistically significant at the 5% level ($p = 0.020$). HMZ employ the 5% threshold as their benchmark for statistical significance throughout the paper. In terms of magnitude, the associated drop in approval ratings is implied to be 4.3

percentage points.

Our analysis first shows this result only emerges after the *inclusion* of five binary variables capturing whether the country experienced a national election in the current or preceding four quarters. Re-estimating that regression without these covariates renders the coefficient of interest statistically insignificant at conventional levels ($p = 0.104$). Nevertheless, one could argue election quarters constitute important confounders that have to be accounted for. To better understand the underlying dynamics, we manually re-coded these election quarter variables and identified a number of incorrectly specified observations that we detail in full in Table A.1. Re-estimating HMZ's regression with the corrected election quarter variables renders the coefficient of interest barely statistically insignificant at the 5% level ($p = 0.058$).

Nevertheless, statistical significance at the 6% level may still constitute an important finding. To systematically explore the role of HMZ's five sets of covariates, we calculate the estimated coefficient of interest for all possible variable combinations, i.e., from $2^5 = 32$ specifications. In these, we employ the corrected election quarter variables. We find none of the 32 specifications yields an estimate that would be statistically significant at the 5% level, while 17 would clear the 10% threshold level.

Next, we inspect HMZ's global panel database and the main variable of interest pertaining to woman executive leadership. Only nine of the 4,328 observations (or 0.2%) feature a woman leader in the quarter after a major transnational terrorist attack. These nine observations consist of three women leaders from two countries: Macapagal Arroyo from the Philippines (three observations), as well as Margaret Thatcher (five observations) and Theresa May from the UK (one observation). As HMZ aim to generalize their hypothesis beyond the May-Manchester observation, we re-estimate their global regression when excluding the May-Manchester observation of the UK in 2017q2. We do so because HMZ's first contribution is to show how May's approval rates indeed suffered after the Manchester attack. Since HMZ's aim is to explore a *general* relationship between leader gender and approval ratings after a major terrorist attack beyond the May-Manchester case, it is worth exploring what happens in the global sample after removing that observation.

Using HMZ's database (that includes the above-mentioned coding errors in election quarters), the corresponding result produces substantially less statistical support for systematic gender differences ($p =$

0.098 with the associated coefficient dropping by 24% from -4.333 to -3.279). This conclusion is further strengthened once we employ the corrected election quarter variables ($p = 0.219$ with a coefficient of -2.390). We then estimate all 2^5 possible combinations of the five groups of control variables while employing the corrected election quarter variables. The results show that none of these specifications produces a coefficient of interest that is statistically significant at the 5% or 10% level.

Similarly, results disappear firmly in terms of statistical relevance once we cluster standard errors at the country level, thereby allowing for a more flexible degree of autocorrelation than HMZ's imposed *AR1* structure ($p = 0.558$ with a coefficient of -3.194 for HMZ's specification). Again re-estimating all 2^5 possible models implies none of the respective estimates would be statistically significant at the 10% level, whether we include or exclude the May-Manchester observation.

Finally, we identified three additional coding errors that, however, remain relatively inconsequential for the statistical robustness of the estimate: (i) the woman executive variable was coded incompletely and partly incorrectly; (ii) Macedonia features duplicate observations from 2006q2 until 2016q2; and (iii) terrorist attacks are mechanically coded as equalling zero for all countries in 1993 – a year in which the *GTD* does not provide any data.

Taken together, we contest the empirical evidence is insufficient to justify drawing systematic links between the gender of a political leader and approval ratings after a transnational terror attack. The data do not support statements such as "women leaders cannot count on rallies following major terrorist attacks", as HMZ write in their abstract.

2 Theresa May Analysis

We first replicate HMZ's analysis of Theresa May's approval rates surrounding the 2017 Manchester Arena attack on May 22, 2017. HMZ first study Wave 12 (conducted before and after the attack in 2017) and then incorporate preceding and subsequent waves in "a difference-in-difference design with fixed effects for time" (HMZ, p.253). In both cases, the derived estimates suggest a firm decrease in approval ratings after the attack ($p < 0.001$) when studying respondents' (i) favorability towards May or (ii) preference for May as best prime minister.

A close exploration of the panel database, however, reveals two coding errors. First, the dependent variable in column (2) of HMZ's Table 2 is coded as one (preferring May as best PM) for observations in which participants did not select May: 35,060 of the 97,155 observations come from survey Waves 7 and 8 in which the survey question referred to then-Prime Minister David Cameron. Second, HMZ state that regression to be a logistic regression – however, the implemented Stata command is *xtreg*, which conducts generalized least squares regressions on panel datasets.

We correct both errors, first independently and then jointly, before re-estimating the affected regression. Results remain statistically robust ($p < 0.001$ in all cases), i.e., our robustness exercises confirm that May's approval ratings indeed suffered a slump after the Manchester attack. Full results are referred to Table B1.

3 Global Analysis

3.1 Data and Methodology

To investigate whether women (compared to men) leaders *generally* suffer decreased approval ratings after a major transnational terror attack, HMZ then assemble a sample of 4,328 observation in 44 countries from 1975 to 2017.[1] Their Table 4 reports results from one regression, predicting executive approval ratings in country i and quarter $t+1$ with

$$\begin{aligned} Approval_{i,t+1} = &\beta_0 + \beta_1 \big(International\ terrorist\ attack\big)_{i,t} \times \big(Woman\ executive\big)_{i,t+1} \\ &+ \mathbf{X}_{i,t}\beta_2 + \lambda_i + \epsilon_{i,t}. \end{aligned} \tag{1}$$

The coefficient β_1 identifies the link between a leader's public approval ratings and whether the country features a woman leader while having experienced a major transnational terror attack in the preceding quarter. HMZ's main definition of a terror attack relates to "any attack that involved an international component (using the definition from the GTD)" that "had more than 15 deaths" (HMZ, p.260). The vector $\mathbf{X}_{i,t}$ incorporates the individual variables of *International terrorist attack* and *Woman executive*, as well as variables measuring "GDP, inflation (logged), the left-center-right placement of the leader, and

[1] On p.250, HMZ incorrectly state they study 66 (rather than 44) countries.

election in that year [specifically, five variables for elections in the current quarter and the preceding four quarters]". Finally, λ_i captures country-fixed effects, and $\epsilon_{i,t}$ constitutes the conventional error term.

HMZ explain that these five sets of covariates are derived from the "executive approval scholarship (Carlin et al., 2020)". Since it lies beyond the scope of our manuscript to theoretically challenge whether that is the appropriate set of covariates, our re-analyses will stay within these sets of regressors, i.e., we do not add any other control variables. For detailed discussions of the intuition underlying the inclusion of each of these variables, we refer to Carlin et al. (2020) and references therein.

In terms of methodology, HMZ employ the Stata command $xtpcse$, designed to estimate a linear regression with panel-corrected standard errors. They impose a panel-specific AR1 autocorrelation structure (command *correlation(psar1)*) and assume independent errors across panels (command *independent*).

3.2 The Role of Covariates

We access HMZ's code and database provided to the *American Political Science Review*. First, we successfully replicate their estimation of equation (1), with the corresponding results displayed in column (1) of Panel A in Table 1 below.[2] Our first order of business is to understand the role of covariates, so we estimate a parsimonious model that only includes the individual variables of *International terrorist attack*, *woman executive*, their interaction term, and country-fixed effects. Column (2) of Panel A documents the corresponding results, yielding a coefficient that is not statistically significant at conventional levels ($p = 0.144$). In terms of magnitude, the coefficient of interest shrinks by almost 40%, from -4.333 to -2.650. Removing country-fixed effects produces similar results ($p = 0.121$; available upon request).

To better understand which control variable produces HMZ's result, we sequentially remove each regressor. That way, we find excluding the election quarter variables suffices to render the coefficient of interest statistically insignificant at conventional levels (see column 3; $p = 0.104$), while reducing its magnitude by 30% (from -4.333 to -2.992). In turn, *only* including the election quarter covariates produces a statistically significant coefficient for the interaction term (see column 4; $p = 0.036$). Taken

[2] Generally, HMZ's results suggest an overall positive rally effect for men leaders (coefficient of 3.112 with a standard error of 0.875).

Table 1: Main results, predicting approval ratings in country i and quarter $t+1$. All regressions include country-fixed effects. Control variables[a] are included unless specified otherwise in the column header.

	Original	No controls	Excl. election quarters	Incl. election quarters only	Incl. corrected election quarters only	Incl. corrected election quarters & controls
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original sample and data						
International terrorist attack$_t$ × woman executive$_t$	-4.333*	-2.650	-2.992	-3.855*	-3.085	-3.452
	(1.859)	(1.812)	(1.841)	(1.834)	(1.797)	(1.821)
Observations	4,328	4,989	4,429	4,835	4,989	4,429
R^2	0.577	0.416	0.458	0.542	0.479	0.531
Panel B: Excluding May-Manchester observation (UK, 2017q2)						
International terrorist attack$_t$ × woman executive$_t$	-3.279	-1.899	-2.224	-2.795	-2.027	-2.390
	(1.982)	(1.947)	(1.970)	(1.963)	(1.926)	(1.945)
Observations	4,327	4,988	4,428	4,834	4,988	4,428
R^2	0.578	0.417	0.459	0.543	0.480	0.533

Notes: Displaying linear regression results, employing the Stata command $xtpcse$. $*p < 0.05$. [a]Includes presence of a woman executive, the binary variable capturing an international terrorist attack, GDP, logged inflation, the left-center-right placement of the leader, and binary variables for elections in that quarter or the previous four quarters.

together, columns (1)-(4) of Panel A illustrate the crucial role of the election quarter variables in producing HMZ's global result.

3.3 Coding Errors: Election Quarter Variables

Next, we closely inspected the election quarter variables, which appear crucial to producing HMZ's result, and identified several objective coding errors. For example, no elections had been coded for the US, while Japanese elections were only coded for one house of the National Diet (Japan's bicameral parliament). In addition, run-off elections and multiple-round elections that spanned several quarters were sometimes coded for only one quarter, although HMZ's coding does not systematically always capture the first or second election. Finally, we identified coding inaccuracies that did not seem to follow a systematic pattern with Austria 2016q4, Ecuador 1997q4, Iceland 2008q2 and 2012q2, Mexico 1991q3, Paraguay 2000q3, and Turkey 2014q3.

To derive an accurate measure of election quarters, we re-coded the variable using publicly available data sources. Table A.1 lists all elections we identified and whether they appear in HMZ's dataset. For the contemporaneous election quarter variable, we update 60 observations in total.

Further, we manually code the election lag variables. This allows us to account for elections that occurred prior to the beginning of the database for a given country that would not be considered otherwise. This accounts for 26 elections across 22 countries, leading to 77 data alterations across 70 observations. When considering the new lags caused by the changes to the contemporaneous variable, we count 317 total amendments across the four lagged variables.

In columns (5) and (6) of Table 1, we put our newly constructed election quarter variables to use. In column (5), we re-estimate column (4), just with the re-coded election quarter variables. Next, column (6) re-estimates HMZ's original regression (our column 1) using the re-coded election quarter variables. In both cases, the coefficient of interest is no longer statistically significant at the 5% level ($p = 0.086$ and $p = 0.058$).

To systematically explore how these corrected election quarter variables affect the universe of possible variable group combinations, we estimate all possible combinations of the five sets of covariates, i.e., $2^5 = 32$ specifications. All of these specifications include the woman executive leadership variable, the binary indicator for having experienced a transnational terrorist attack, and the interaction term of the two. In the absence of clear theoretical guidance on the 'correct' set of control variables, these robustness checks can provide valuable information about the stability of the coefficient of interest. *Panel A* of Figure 1 visualizes the corresponding results for the coefficient of interest associated with the interaction term. Overall, none of the estimated coefficients would cross the 5% hurdle of statistical significance, but 17 of them would be statistically significant at the 10% level.

3.4 The May-Manchester Case

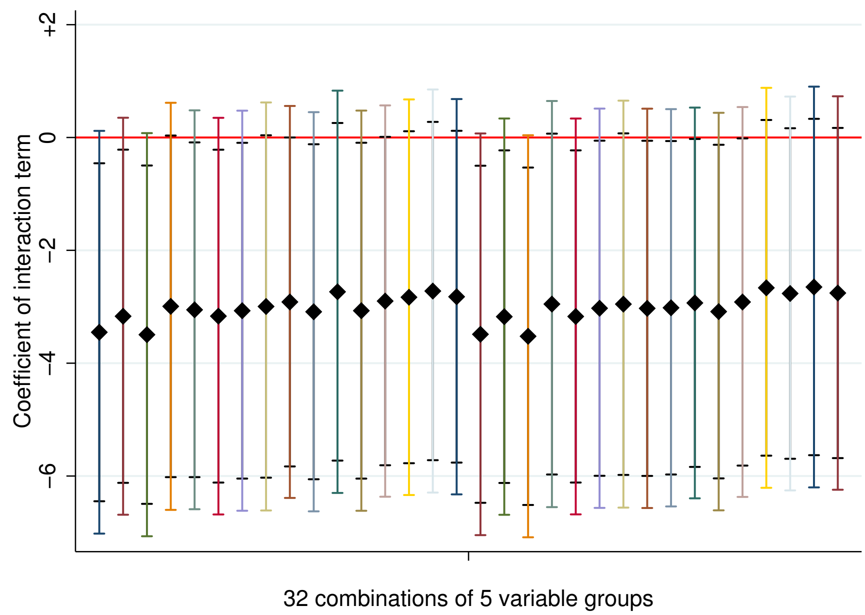
Statistical significance at the 10% level may still constitute an interesting result, potentially hinting at a meaningful underlying pattern. To better understand the generality of that finding, we next focus on the characteristics of the interaction term variable, i.e., the cases in which a country is led by a woman executive and has just experienced a major transnational terrorist attack. Only nine (or 0.02%) of the 4,328 observations meet those criteria.[3] Table 2 summarizes these cases that only pertain to three leaders in two countries: Gloria Macapagal Arroyo in the Philippines; Margaret Thatcher and Theresa May in the UK. Three of these observations immediately follow another observation with the Philippines experiencing terrorist attacks in three consecutive quarters (2002q4, 2003q1, and 2003q2) and the UK experiencing such tragedies in 1988q3 and 1988q4.

Table 2 also documents changes in approval ratings compared to the previous quarters (that were also characterized by the same leaders). To provide background, the average quarter-to-quarter change in approval ratings in HMZ's sample is -0.05ppt (standard deviation of 5.5ppt). While Macapagal Arroyo did suffer a drop of 10.87ppt in 2003q4 (which equates to almost two standard deviations), she also gained 12.01ppt in the subsequent quarter that also followed a transnational terrorist attack. Considering the UK cases, all of Margaret Thatcher's changes in approval ratings appear relatively minor in magnitude, reaching an absolute maximum of -2.36ppt after the 1988q4 attack.

Finally, the May-Manchester observation is indeed standing out with a drop of 11.47ppt, equivalent to more than two standard deviations. To better understand whether that case truly generalizes, we

[3] Implementing a lower threshold for an international terrorist attack with 10+ (rather than 16+) deaths expands that number from nine to 16. All our results are consistent when employing that threshold (see Figure B1).

Panel A: Employing corrected election quarter variables



Panel B: Employing corrected election quarter variables and excluding May-Manchester observation

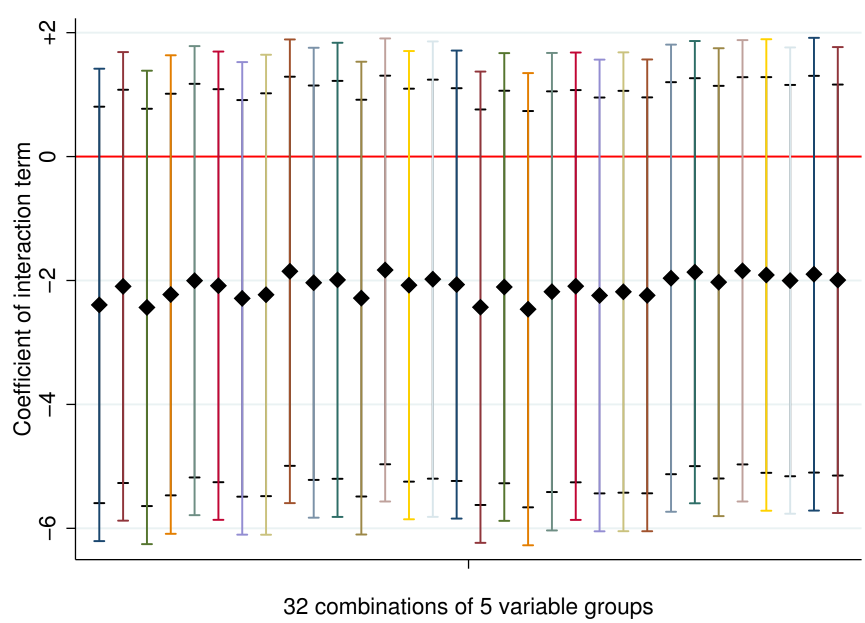


Figure 1: Displaying the estimated coefficient of interest (interaction term between terrorist attack and woman leader) from 32 regressions featuring all 2^5 possible combinations of control variable groups. In both graphs, the first coefficient on the left comes from the full specification (as in HMZ). 95% and 90% confidence intervals are displayed.

Table 2: Women leaders after major transnational terrorist attacks.

Country	Quarter	Leader	Approval this quarter	Approval next quarter	Change in ratings
Philippines	2002q4	Macapagal Arroyo	36.39%	25.52%	-10.87ppt
Philippines	2003q1	Macapagal Arroyo	25.52%	37.53%	+12.01ppt
Philippines	2003q2	Macapagal Arroyo	37.53%	35.14%	-2.39ppt
UK	1979q3	Thatcher	36.34%	35.94%	-0.40ppt
UK	1982q4	Thatcher	38.96%	40.12%	+1.16ppt
UK	1985q1	Thatcher	34.24%	32.70%	-1.54ppt
UK	1988q3	Thatcher	39.34%	39.08%	-0.26ppt
UK	1988q4	Thatcher	39.08%	36.72%	-2.36ppt
UK	2017q2	May	43.51%	32.04%	-11.47ppt

re-run HMZ's regression after excluding just the May-Manchester observation from 2017q2. Panel B of Table 1 reports the derived estimates. In this case, HMZ's specification (that features incorrectly coded election quarter variables) produces a coefficient that drops well below the 5% threshold level of statistical relevance ($p = 0.098$). None of the estimates in Panel B becomes statistically significant at conventional levels, and re-estimating HMZ's specification with the correctly coded election quarter variables in column (6) produces a p-value of 0.219.

Panel B of Figure 1 visualizes the 32 estimates from implementing all possible combinations of the five groups of covariates, and none of these cross the 10% threshold level of statistical significance. In sum, the May-Manchester case carries an outsized impact on the statistical significance of the global finding, suggesting there may be a factor unique to Theresa May in 2017q2 that drives HMZ's result. Section 4 will expand on that possibility.

3.5 Additional Coding Errors

Beyond the coding errors pertaining to election quarters, we identified three sampling errors that turned out to be inconsequential to our conclusions. First, the woman executive variable had 376 missing observations. We reconstruct that variable from publicly available data and further identify coding errors in observations that did not have missing values. Due to the use of time period $t + 1$ in the approval variable, and to stay as consistent with HMZ as possible, we define the variable by any quarter in which a woman held the office of head of government or state for at least 15 days in the corresponding quarter. In total, we adjust 612 observations.

Second, HMZ's database features duplicate observations for Macedonia from 2006q2 to 2016q2. While other countries also feature multiple entries in the original Executive Approval Dataset (Carlin et al., 2019), HMZ generally prefer entries using the 'EXEC' suffix (examining a summary measure of all available approval measures for systems with direct and indirect elections for different heads of state) where possible. This exists in Macedonia's case, so we omit the duplicate Macedonian observations.

Third, HMZ access the GTD for data on international terrorist attacks throughout their sample period. However, the GTD does not feature any data for 1993, as that information was lost in an office move (LaFree and Dugan, 2007; START, 2022). HMZ mechanically code all terrorist attack observations in 1993 with a zero.

Appendix Figure B2 visualizes the corresponding estimates after implementing all three adjustments, where we again estimate all 32 possible models. *Panel A* employs the full sample, while *Panel B* excludes the May-Manchester observation. The corresponding results are comparable to those from Figure 1, implying these additional coding errors remain inconsequential to the result of interest.

3.6 The Calculation of Standard Errors

Finally, we explore the issue of calculating standard errors. HMZ implement the Stata command $xtpcse$ that is designed to estimate a linear regression with panel-corrected standard errors and an autocorrelation of an order of one. In addition, their regression command calculates independent errors across panels, i.e., assuming there to be one disturbance that is common to all observations.

While we do not suggest that specification to be necessarily incorrect, we explore what happens to the coefficient of interest once we employ a more flexible approach when it comes to both the autocorrelation structure and the behavior of standard errors across countries. To do so, we deploy the $xtreg$ command and cluster standard errors at the country level, which allows for flexible degrees of autocorrelation within countries, recognizing cross-country differences in the behavior of the error term. Re-estimating

HMZ's exact regression yields a coefficient that is statistically insignificant at any conventional levels ($p = 0.558$). The corresponding results from all 32 specifications, visualized in Figure B3, confirm that none of the 32 specifications yields an estimate that would approach the 10% threshold level of statistical significance. This is also the case once we remove the May-Manchester observation.

4 Discussion

How can we interpret HMZ's regression result in light of these robustness checks? In the absence of a clear theoretical framework to suggest the 'correct' empirical specification, researchers often consider a combination of several alternative specifications as collective evidence. Taken together, HMZ's global result does not pass the majority of alternatives we suggest here, however, which means we should be careful to generalize the robust observation of Theresa May's approval ratings dropping significantly after the Manchester arena attack. It is important, though, to theorize how else we could reconcile the May results with the fragility of the global estimate. We want to offer four such avenues. While that list is not meant to be complete, we hope to illustrate some of the more plausible explanations.

4.1 Gender vs. Other Characteristics

First, characteristics other than gender may explain the May-Manchester result. In fact, the May-Manchester setting (May 22, 2017) comes during a tumultuous period preceding a general election (Jun 8, 2017) and just a year after the UK's historic Brexit vote (June 23, 2016). Perhaps the UK populace indeed felt differently about May than they would have for another prime minister after the Manchester attack – but gender need not necessarily be that delineating aspect.

4.2 Leader Gender vs. Leader Response

Second, it remains difficult in the global framework to separate out the public's reaction to the terror attack from the public's reaction to *the leader's response* to the attack. For example, Frey et al. (2007) summarize "it might be difficult to isolate the effects of terrorism from those caused by government reactions." Of course, this holds true for women and men leaders alike; but it introduces substantial measurement error into equation (1).

For instance, Jacinda Ardern's measured response to the Christchurch mosque shootings in 2019 has been lauded nationally and internationally, and anecdotal evidence suggests she enjoyed a rise in popularity because of that response (Fifield, 2019; Maya Salam, 2019). Another prominent case relates to George Bush's response to 9/11, appealing to US Americans' patriotism and unity, and offering immediate (although controversial) policy responses. Finally, one may also think about the 2004 Madrid train bombings that have been linked to the Conservative party losing national elections thereafter (e.g., see Montalvo, 2011): In that case, the political leader (José María Aznar) blamed another group that turned out to be unconnected to the attack – a reaction that has not been perceived favorably by the electorate (e.g., see BBC News, 2004). Concerning Theresa May's response after Manchester, some of her statements and actions may have alienated the electorate. For instance, she declared "if human rights laws stop us from doing it [restricting the freedom and movements of terrorist suspects], we will change those laws so we can do it" (Mason and Dodd, 2017, cited in Lenard, 2018).

4.3 Reverse Causality

Third, reverse causality could influence the derived estimates, even though HMZ predict approval ratings in the quarter after the corresponding attack. Studying centuries of European data, Dube and Harish (2020) show "polities led by queens engaged in war more than polities led by kings." Their analysis further suggests "single queens were more likely to be attacked than single kings". While we do not know whether such dynamics translate to the 20^{th} and 21^{st} centuries of HMZ's sample, terror groups may explicitly target polities led by women (but not men) they perceive to be 'weak', which may correlate with a subsequent drop in approval ratings.

4.4 Omitted Variables

Fourth, omitted variables are always a concern, especially when dealing with data aggregated at the country level and comparing a variety of nations that differ in size, political, religious, and societal structures, as well as other characteristics that are potentially relevant to the key variables considered. We can test how influential omitted variables would need to be in order to nullify a statistically significant result. Specifically, Cinelli and Hazlett's (2020) test (Stata command *sensemakr*) allows identifying the total amount of residual variance in both treatment and outcome variables that an omitted variable would need to explain in order to reduce a point estimate to zero. In HMZ's case, that test reveals omitted variables explaining only 1.14% of the total variance would be sufficient to yield a point estimate of the interaction term that equals zero. Thus, HMZ's finding is particularly vulnerable to omitted variable bias beyond the specifications discussed above.

4.5 Sample Size

Finally, it is of course possible that women leaders are indeed perceived differently from men leaders after a transnational terrorist attack – maybe just as HMZ hypothesize. As we know, absence of evidence does not necessarily constitute evidence of absence. As highlighted in Section 3.4, HMZs global analysis is only able to include nine observations in which a woman leader experienced a transnational terrorist attack. While there is no objective threshold of the number of observations below which a statistical analysis becomes uninformative (at least to our knowledge), it is possible that we simply do not have enough data to properly test HMZ's hypothesis yet. If that was the case, however, our re-analyses heed caution not to overstate the empirical evidence at hand.

5 Conclusion

HMZ's analysis of the Theresa May case uses comprehensive data and a strong identification strategy to convincingly show May's approval ratings suffered following the Manchester attack. However, our replication of their global analysis leads us to conclude that the empirical evidence does not justify the general claim made in the final sentence of HMZ's abstract that "conventional theory on rally events requires revision: women leaders cannot count on rallies following major terrorist attacks" (p. 249).

References

BBC News (2004). Scores die in Madrid bomb carnage.

Carlin, R. E., Carreras, M., and Love, G. J. (2020). Presidents' sex and popularity: Baselines, dynamics and policy performance. *British Journal of Political Science*, 50(4):1359–1379.

Carlin, R. E., Hartlyn, J., Hellwig, T., Love, G. J., Martinez-Gallardo, C., and Singer, M. M. (2019). Executive Approval Database 2.0.

Cinelli, C. and Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1):39–67.

Dube, O. and Harish, S. (2020). Queens. *Journal of Political Economy*, 128(7):2579–2652.

Fifield, A. (2019). New Zealand's prime minister receives worldwide praise for her response to the mosque shootings.

Frey, B. S., Luechinger, S., and Stutzer, A. (2007). Calculating tragedy: Assessing the costs of terrorism. *Journal of Economic Surveys*, 21(1):1–24.

Holman, M. R., Merolla, J. L., and Zechmeister, E. J. (2022). The curious case of Theresa May and the public that did not rally: Gendered reactions to terrorist attacks can cause slumps not bumps. *American Political Science Review*, 116(1):249–264.

LaFree, G. and Dugan, L. (2007). Introducing the Global Terrorism Database. *Terrorism and Political Violence*, 19(2):181–204.

Lenard, P. T. (2018). Democratic citizenship and denationalization. *American Political Science Review*, 112(1):99–111.

Mason, R. and Dodd, V. (2017). May: I'll rip up human rights laws that impede new terror legislation.

Maya Salam (2019). Jacinda Ardern Is Leading by Following No One.

Montalvo, J. G. (2011). Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections. *Review of Economics and Statistics*, 93(4):1146–1154.

START (2022). Global Terrorism Database Data.

Online Appendix A: Data Documentation

Table A.1: Comparing election coding in original HMZ dataset to our dataset (1/5).

Country	Elections										Comments
Argentina	1983q4	1985q4	1987q3	1989q2	1991q3	1991q4	1993q4	1995q2	1997q4	1999q4	Includes both general and
In dataset?	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	midterm elections.
Argentina	2001q4	2003q2	2005q4	2007q4	2009q2	2011q4	2013q4	2015q4	2017q4		1991 election spanned 4 months.
In dataset	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Australia	1975q4	1977q4	1980q4	1983q1	1984q4	1987q3	1990q1	1993q1	1996q1	1998q4	Election in 1974q2.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1975q1.
Australia	2001q4	2004q4	2007q4	2010q3	2013q4	2016q3					
In dataset?	✓	✓	✓	✓	✓	✓					
Austria	1990q4	1992q2	1994q4	1995q4	1998q2	1999q4	2002q4	2004q2	2006q4	2008q3	Election in 1986q4 - dataset starts 1987q2.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Includes legislative and presidential
Austria	2010q2	2013q3	2016q2	2016q4	2017q4						elections. 2016q2 results were annulled,
In dataset?	✓	✓	✓	✗	✓						election reconducted in q4
Bolivia	2002q2	2005q4	2009q4	2014q4							None.
In dataset?	✓	✓	✓	✓							
Brazil	1982q4	1986q4	1989q4	1990q4	1994q4	1998q4	2002q4	2006q4	2010q4	2014q4	None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Bulgaria	1990q2	1991q4	1992q1	1994q4	1996q4	1997q2	2001q2	2001q4	2005q2	2006q4	Includes parliamentary and presidential
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	elections.
Bulgaria	2009q3	2011q4	2013q2	2014q4	2016q4	2017q1					
In dataset?	✓	✓	✓	✓	✗	✓					
Canada	1979q2	1980q1	1984q3	1988q4	1993q4	1997q2	2000q4	2004q2	2006q1	2008q4	None
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Canada	2011q2	2015q4									
In dataset?	✓	✓									
Chile	1993q4	1997q4	1999q4	2000q1	2001q4	2005q4	2006q1	2009q4	2010q1	2013q4	Chilean electoral structure has elections
In dataset?	✓	✓	✓	✗	✓	✓	✗	✓	✗	✓	in Q4, with run-offs in following Q1
Chile	2017q4										Years (1999, 2005, 2009)
In dataset?	✓										

Notes: Checkmark indicates election appeared in HMZ's original dataset. Cross indicates election did not feature in HMZ's original dataset.

Table A.1 continued: Comparing election coding in original HMZ dataset to our dataset (2/5).

Country	Elections										Comments
Colombia	1994q1	1994q2	1998q1	1998q2	2002q1	2002q2	2006q1	2006q2	2010q1	2010q2	Includes parliamentary and
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	presidential elections.
Colombia	2014q1	2014q2									
In dataset?	✓	✓									
Costa Rica	1982q1	1986q1	1990q1	1994q1	1998q1	2002q1	2002q2	2006q1	2010q1	2014q1	2002 and 2014 had second rounds.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Czech Republic	1990q2	1992q2	1996q2	1998q2	2002q2	2006q2	2010q2	2013q4	2017q4		1990 and 1992 as Czechoslovakia
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Denmark	1975q1	1977q1	1979q4	1981q4	1984q1	1987q3	1988q2	1990q4	1994q3	1998q1	Denmark's data stops in 2014q4
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Denmark	2001q4	2005q1	2007q4	2011q3							
In dataset?	✓	✓	✓	✓							
Dominican Republic	2002q2	2004q2	2006q2	2008q2	2010q2	2012q2	2016q2				None.
In dataset?	✓	✓	✓	✓	✓	✓	✓				
Ecuador	1984q1	1984q2	1986q2	1988q1	1988q2	1990q2	1992q2	1992q3	1994q2	1996q2	Multiple round elections (84,88,
In dataset?	✓	✗	✓	✓	✗	✓	✓	✗	✓	✓	92,96,98,17).
Ecuador	1996q3	1997q4	1998q2	1998q3	2002q4	2006q4	2009q2	2013q1	2017q1	2017q2	
In dataset?	✗	✗	✓	✗	✓	✓	✓	✓	✓	✓	
El Salvador	1988q1	1989q1	1991q1	1994q1	1994q2	1997q1	1999q1	2000q1	2003q1	2004q1	Parliamentary and presidential,
In dataset?	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	some run-offs.
El Salvador	2006q1	2009q1	2012q1	2014q1	2015q1						
In dataset?	✓	✓	✓	✓	✓						
France	1978q1	1981q2	1986q1	1988q2	1993q1	1995q2	1997q2	2002q2	2007q2	2012q2	Parliamentary and presidential.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
France	2017q2										
In dataset?	✓										
Germany	1976q4	1980q4	1983q1	1987q1	1990q4	1994q4	1998q3	2002q3	2005q3	2009q3	East German elections excluded.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	All elections prior to 1990 pertain
Germany	2013q3	2017q3									to West Germany.
In dataset?	✓	✓									

Notes: Checkmark indicates election appeared in HMZ's original dataset. Cross indicates election did not feature in HMZ's original dataset.

Table A.1 continued: Comparing election coding in original HMZ dataset to our dataset (3/5).

Country	Elections										Comments
Greece	1989q2	1989q4	1990q2	1993q4	1996q3	2000q2	2004q1	2007q3	2009q4	2012q2	None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Greece	2015q1	2015q3									
In dataset?	✓	✓									
Guatemala	1990q4	1991q1	1994q3	1995q4	1996q1	1999q4	2003q4	2007q3	2007q4	2011q3	Multiple round elections.
In dataset?	✓	✗	✓	✓	✗	✓	✓	✓	✗	✓	
Guatemala	2011q4	2015q3	2015q4								
In dataset?	✗	✓	✗								
Honduras	1989q4	1993q4	1997q4	2001q4	2005q4	2009q4	2013q4	2017q4			None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓			
Hungary	2002q2	2006q2	2010q2	2014q2							There was an election in 1998q2.
In dataset?	✓	✓	✓	✓							Dataset starts 1998q3.
Iceland	1987q2	1988q2	1991q2	1995q2	1996q2	1999q2	2003q2	2004q2	2007q2	2009q2	Contains presidential and parliamentary
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	elections. Some presidential elections
Iceland	2012q2	2013q2	2016q2	2016q4	2017q4						uncontested (1984, 1992, 2000, 2008).
In dataset?	✗	✓	✓	✓	✓						2008 incorrectly coded in original dataset.
Ireland	1981q2	1982q1	1982q4	1987q1	1989q2	1990q4	1992q4	1997q2	1997q4	2002q2	None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Ireland	2007q2	2011q1	2011q4	2016q1							
In dataset?	✓	✓	✓	✓							
Italy	1992q2	1994q1	1996q2	2001q2	2006q2	2008q2	2013q1				None.
In dataset?	✓	✓	✓	✓	✓	✓	✓				
Japan	1976q4	1977q3	1979q4	1980q2	1983q2	1983q4	1986q3	1989q3	1990q1	1992q3	Includes both houses of parliament.
In dataset?	✓	✗	✓	✓	✗	✓	✓	✗	✓	✗	
Japan	1993q3	1995q3	1996q4	1998q3	2000q2	2001q3	2003q4	2004q3	2005q3	2007q3	
In dataset?	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	
Japan	2009q3	2010q3	2012q4	2013q3	2014q4	2016q3	2017q4				
In dataset?	✓	✗	✓	✗	✓	✗	✓				
Macedonia	2006q3	2008q2	2009q1	2009q2	2011q2	2014q2					Macedonia dataset ends 2016q2.
In dataset?	✓	✓	✓	✗	✓	✓					2009 had run-off.

Notes: Checkmark indicates election appeared in HMZ's original dataset. Cross indicates election did not feature in HMZ's original dataset.

Table A.1 continued: Comparing election coding in original HMZ dataset to our dataset (4/5).

Country	Elections										Comments
Mexico	1991q3	1994q3	1997q3	2000q3	2003q3	2006q3	2009q3	2012q3	2015q2		General election 1988q3.
In dataset?	✗	✓	✓	✓	✓	✓	✓	✓	✓		Dataset starts 1989q1.
New Zealand	1990q4	1993q4	1996q4	1999q4	2002q3	2005q3	2008q4	2011q4	2014q3	2017q3	None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Nicaragua	1990q1	1996q4	2001q4	2006q4	2011q4	2016q4					None.
In dataset?	✓	✓	✓	✓	✓	✓					
Panama	1994q2	1999q2	2004q2	2009q2	2014q2						None.
In dataset?	✓	✓	✓	✓	✓						
Paraguay	1998q2	2000q3	2003q2	2008q2	2013q2						2000 was vice-presidential election.
In dataset?	✓	✗	✓	✓	✓						
Peru	1985q2	1990q2	1992q4	1995q2	2000q2	2001q2	2006q2	2011q2	2016q2		Election in 1980q2.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓		Dataset starts 1981q1.
Philippines	1987q2	1992q2	1995q2	1998q2	2001q2	2004q2	2007q2	2010q2	2013q2	2016q2	Presidential election in 1986q1.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1986q2.
Poland	1990q4	1991q4	1993q3	1995q4	1997q3	2000q4	2001q3	2005q3	2005q4	2007q4	Legislative election in 1989q2.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1989q3.
Poland	2010q2	2010q3	2011q4	2015q2	2015q4						
In dataset?	✓	✗	✓	✓	✓						
Portugal	1987q3	1991q1	1991q4	1995q4	1996q1	1999q4	2001q1	2002q1	2005q1	2006q1	Elections in 1985q4 and 1986q1.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1986q2.
Portugal	2009q3	2011q1	2011q2	2015q4	2016q1						
In dataset?	✓	✓	✓	✓	✓						
Russia	1996q3	1999q4	2000q1	2003q4	2004q1	2007q4	2008q1	2011q4	2012q1	2016q3	Legislative election in 1995q4.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1996q3.
South Korea	1988q2	1992q1	1992q4	1996q2	1997q4	2000q2	2002q4	2004q2	2007q4	2008q2	Presidential election in 1987q4.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1988q2.
South Korea	2012q2	2012q4	2016q2	2017q2							
In dataset?	✓	✓	✓	✓							
Spain	1986q2	1989q4	1993q2	1996q1	2000q1	2004q1	2008q1	2011q4	2015q4	2016q2	None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

Notes: Checkmark indicates election appeared in HMZ's original dataset. Cross indicates election did not feature in HMZ's original dataset.

Table A.1 continued: Comparing election coding in original HMZ dataset to our dataset (5/5).

Country	Elections										Comments
Turkey	2002q4	2007q2	2007q3	2011q2	2014q3	2015q2	2015q4				2007q2 election was annulled.
In dataset?	✓	✗	✓	✓	✗	✓	✓				
Ukraine	2002q1	2004q4	2006q1	2007q3	2010q1	2012q4	2014q2	2014q4			Elections held 1999q4.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓			Dataset starts 2000q2.
United Kingdom	1979q2	1983q2	1987q2	1992q2	1997q2	2001q2	2005q2	2010q2	2015q2	2017q2	Elections held in 1974q4.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Dataset starts 1975q1
United States	1976q4	1978q4	1980q4	1982q4	1984q4	1986q4	1988q4	1990q4	1992q4	1994q4	Election in 1974q4.
In dataset?	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	Dataset starts 1975q1.
United States	1996q4	1998q4	2000q4	2002q4	2004q4	2006q4	2008q4	2010q4	2012q4	2014q4	
In dataset?	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	
United States	2016q4										
In dataset?	✗										
Uruguay	1989q4	1994q4	1999q4	2004q4	2009q4	2014q4					None.
In dataset?	✓	✓	✓	✓	✓	✓					
Venezuela	1988q4	1993q4	1998q4	2000q3	2005q4	2006q4	2010q3	2012q4	2013q2	2015q4	None.
In dataset?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

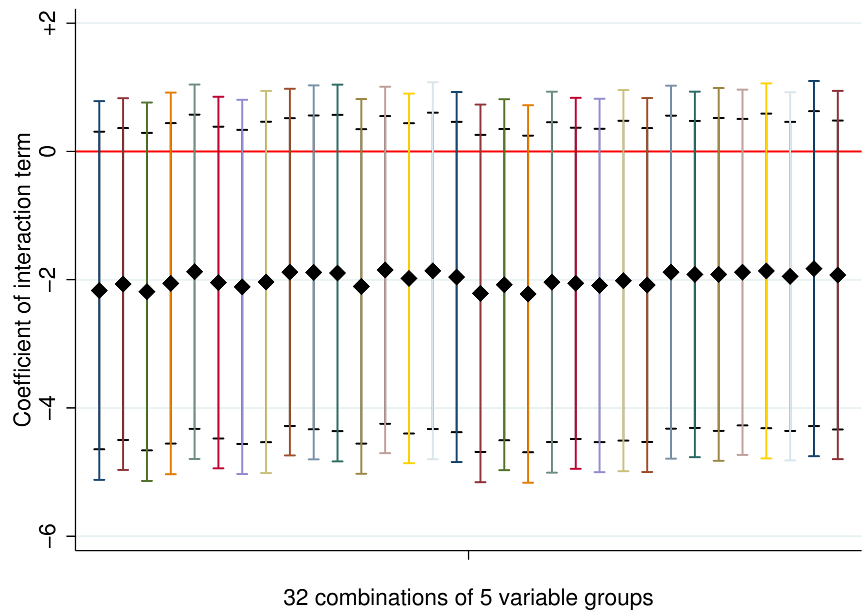
Notes: Checkmark indicates election appeared in HMZ's original dataset. Cross indicates election did not feature in HMZ's original dataset.

Online Appendix B: Additional Specifications

Table B1: Replicating and extending HMZ's Table 2 (entitled "Difference-in-Difference, With Fixed Effects"). Columns (1) and (2) replicate Table 2's columns (1) and (2), while columns (3)-(5) re-estimate column (2), first excluding pre-Wave 9 information, then employing a logit regression, and finally implementing both variations simultaneously.

	(1)	(2)	(3)	(4)	(5)
	May likeability	May best PM			
Manchester attack × time	-0.1274*	-0.0501*	-0.0502*	-0.6806*	-1.0181*
	(0.0187)	(0.0044)	(0.0038)	(0.0576)	(0.0797)
Manchester attack	0.0178	-0.0002	0.0024	0.0198	0.0449
	(0.0278)	(0.0040)	(0.0046)	(0.0527)	(0.0866)
Time	-0.5007*	0.2098*	0.0287*	2.7819*	0.5882*
	(0.0229)	(0.0036)	(0.0029)	(0.0514)	(0.0610)
Constant	1.9443*	0.1505*	0.1999*	-5.3048*	-5.9168*
	(0.0879)	(0.0132)	(0.0149)	(0.1840)	(0.2863)
Controls	✓	✓	✓	✓	✓
Wave-fixed effects	✓	✓	✓	✓	✓
Omitting Waves 1-8?			✓		✓
Logistic regression?				✓	✓
Observations	143,499	97,155	62,095	97,155	62,095

Notes: Standard errors in parentheses. Dependent variables are 11-point favorability scale in column (1) and perceptions of May as the best PM in columns (2)-(5). Full controls include whether someone identifies as ethnically British, gender, Labour party membership, other party membership, income, and ideology. *$p < 0.05$.



Excluding May-Manchester observation

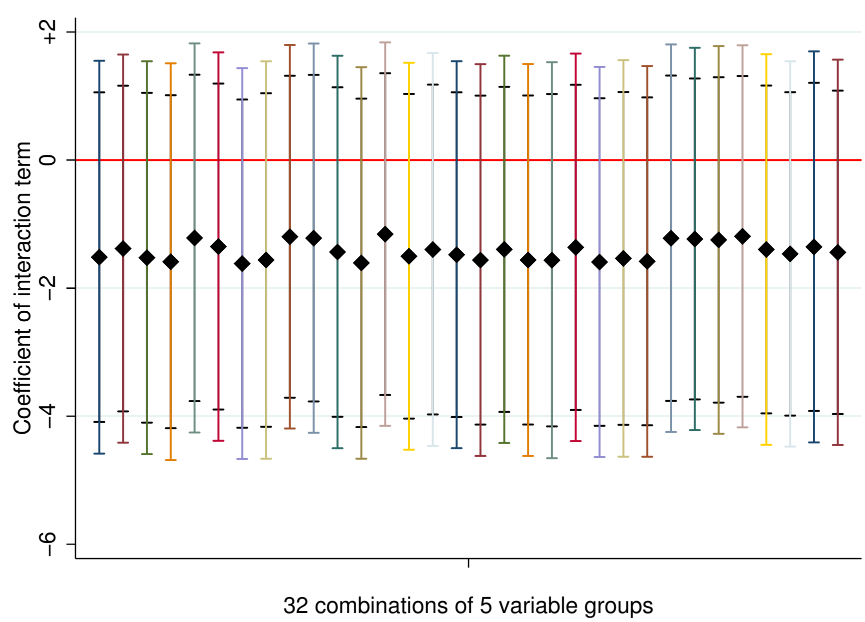
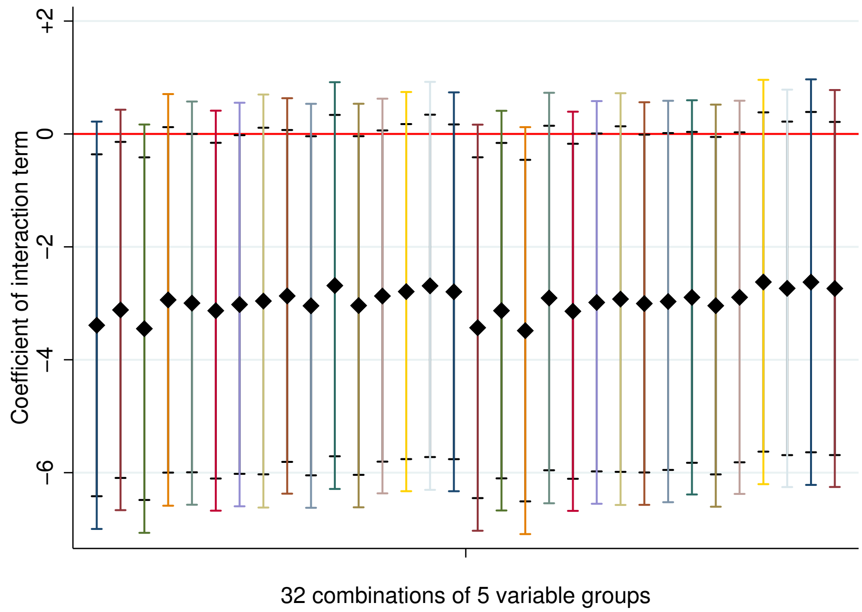


Figure B1: Defining a terrorist attack as leading to 10 or more deaths. Displaying the estimated coefficient of interest (interaction term between terrorist attack and woman leader) for from 32 regressions including all 2^5 combinations of control variable groups. In both graphs, the first coefficient on the left comes from the full specification (as in HMZ). 95% and 90% confidence intervals are displayed.



Excluding May-Manchester observation

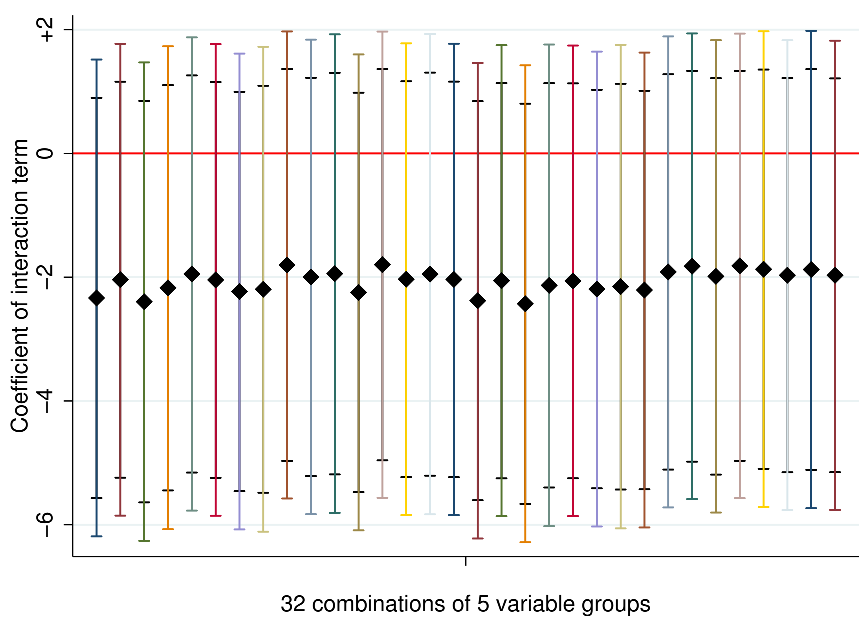


Figure B2: Resolving additional coding errors pertaining to the woman executive variable, Macedonia, and the year 1993 (see Section 3.5). Displaying the estimated coefficient of interest (interaction term between terrorist attack and woman leader) for from 32 regressions including all 2^5 combinations of control variable groups. In both graphs, the first coefficient on the left comes from the full specification (as in HMZ). 95% and 90% confidence intervals are displayed.

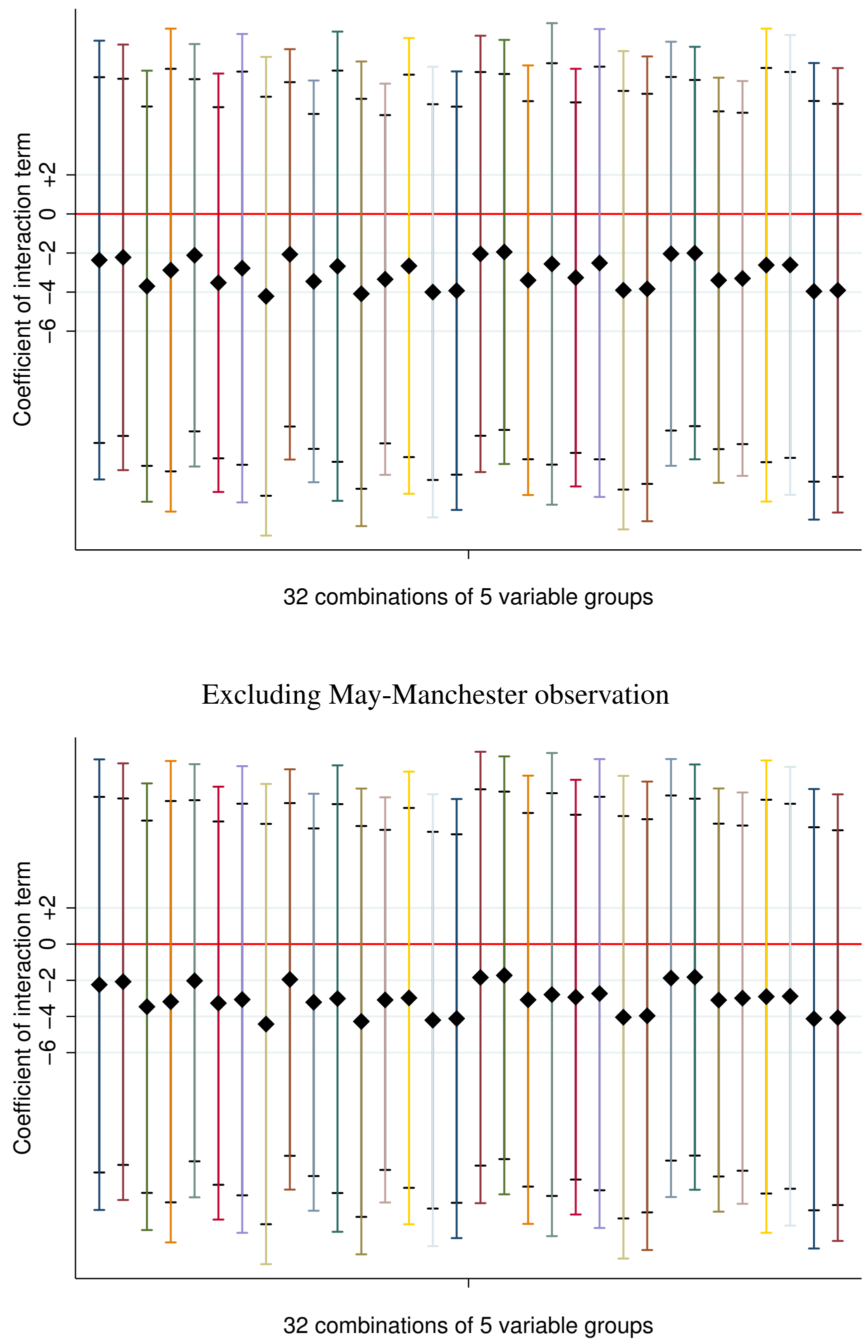


Figure B3: Estimating standard errors clustered at the country level. Displaying the estimated coefficient of interest (interaction term between terrorist attack and woman leader) for from 32 regressions including all 2^5 combinations of control variable groups. In both graphs, the first coefficient on the left comes from the full specification (as in HMZ). 95% and 90% confidence intervals are displayed.