Online Appendix for Media Coverage, Public Interest, and Support in the 2016 Republican Invisible Primary

May 24, 2018

Contents

1 Outline of Appendix

In this appendix, we first outline the Bayesian Structure Vector Autoregression (BSVAR) approach we take in this article. This includes a discussion of what BSVAR is, its benefits and drawbacks, and then the specifics for the implementation in this article. We next present a series of robustness tests that allow us to test assumptions about identification, priors, and how we aggregated the measure of support.

2 Bayesian Structural Vector Autoregression BSVAR

Analysis of media coverage, polling, and public interest presents a challenge. The relationships between all three are expected to be endogenous, with each series responding to changes in the other two. Failure to address such endogeneity issues in the modeling strategy would result in unreliable hypothesis testing (Brandt and Williams 2007). The traditional approach uses Vector Autoregression (VAR) to accommodate endogenous relationships.[1] VAR allows the researcher to estimate the influence of a set of variables on one another by specifying a system of equations which includes lags of the dependent variable as independent variables.

Although VAR analysis is commonly used, it does have limitations. Frequentist VAR requires the identification of individual series dynamics prior to hypothesis testing. In particular, for accurate hypothesis testing the order of integration of each series must be identified. Once all series are identified, the VAR equations that relate them ought to be balanced such that all series have the same order. A variety of tests can be used to check series dynamics, but all of them involve "knife-edge" decisions using a p-value of 0.05, leading to somewhat arbitrary decisions about the dynamics of each series (Brandt and Williams 2007; Williams 1993). This is especially obfuscating when different tests lead to conflicting conclusions, as is the case for our analysis. In addition, tests of the same series for different candidates, which we would expect to follow similar dynamics, lead to conflicting results.[2] Any misidentification during the pre-testing process presents a threat to inference, which is a significant drawback to traditional VAR.

Recently, Bayesian Structural VAR has been proposed as a solution to the limitations of traditional VAR (Brandt and Freeman 2006, 2009).[3] Unlike traditional VAR, Bayesian VAR does not require pretesting for accurate inferences (Sims and Zha 1998). Instead of pre-testing for the dynamics of series or the lag lengths, "prior beliefs" are set that reflect expectation about series dynamics. In Bayesian analysis, priors are a way to quantify preexisting beliefs about the world given previous research. Priors influence the model posterior, but the effects of the prior can be directly investigated by re-estimating relationships using a variety of priors (Jackman 2009).

2.1 Identification of Contemporaneous Shocks

It is possible for one series to affect another immediately or for the effect to play out over several days after the initial change. In the latter case, the lagged values of one series affect the present value of the series of interest. Both of these relationships can be modeled using Bayesian Structural VAR, but the immediate, or contemporaneous, effects require special

[1]Structural Equation Modeling (SEM) can also be used, but requires very strong assumptions about the underlying data generating process. Because of this requirement, SEM is less often used than VAR (Box-Steffensmeier et al. 2014).

[2]Series stationarity was evaluated using Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests.

[3]We investigated several other methods of analysis that could incorporate an unknown degree of stationarity. For example, Fully Modified VAR (FM-VAR) has been suggested as a means of incorporating series with mixed stationarity, but provides only a conservative upper bound for hypothesis testing (Phillips 1995). Additionally, the IRF dynamics using FM-VAR have asymmetric distributions (Phillips 1998; Mills 1998). Toda and Yamamoto (1995) propose a lag-augmented VAR (LA-VAR) approach. Recently, the LA-VAR IRFs have been shown to be relatively robust (Ashley and Verbrugge 2009), but the LA-VAR process still requires pretesting. Finally, there might be some concerns over the bounded nature of some of our series (Grant and Lebo 2016). A Fractionally-Integrated approach might be one way to deal with a bounded series but has its own limitations (Keele, Linn and Webb 2016*a*,*b*).

Table 1: Contemporaneous Identification
Table 1: Model Identification

		Media	Interest	Polling
Media	=	X		
Interest	=	X	X	
Polling	=	X	X	X

Table 1: This table displays the identification of the contemporaneous shocks. Each row represents an equation with the label being the dependent variable. The columns indicate the contemporaneous variable. An X means that the coefficient was estimated, while a blank means that the coefficient was constrained to 0. The diagonals are never constrained to zero by definition because they are used to include the dependent variable. See text for details.

care. It is not possible to identify all contemporaneous relationships simultaneously; in order for a model with p series to be identified, only $p(p+1)/2$ coefficients can be estimated, while the remaining relationships must be constrained to 0 (Brandt and Freeman 2009). Since all of our systems have 3 series, this implies that we can estimate, at most, 6 contemporaneous relationships.

Identification of contemporaneous shocks should be driven by theory and the effects of the identification should be tested. Our hypotheses, although they speak to the relationships between variables, do not make strong predictions about how quickly the effects play out. We impose a single identification that allows for the maximum number of contemporaneous relationships and then test to see if alternative identifications affect our findings.

Our identification is in Table 1. We allow media to have an immediate effect on interest and polling and for interest to have an immediate effect on polling. This is based on the expectation that media and interest influence each other rapidly, but that the effects of polling will take longer to influence the series. In addition, we also estimated our models using two additional identifications: one where only the media could have immediate effects and one where only interest could have immediate effects. Our results are substantively the same and they can be found in Section 4.

2.2 Estimation

We estimate one model for each candidate in our sample.[4] These models focus on a single unit over time, meaning that each candidate is modeled separately. By estimating each candidate separately, we (1) avoid restrictive assumptions that effects must be identical between candidates, (2) prevent any one candidate from driving the results and (3) avoid having to reparameterize to accommodate the fact that our series sum to unity across candidates. Our model still allows for indirect interactions between candidates. For example, an increase in media coverage for Ted Cruz cannot, on its own, decrease Marco Rubio's support in opinion polls. Increased coverage of Cruz may, however, take up space that would otherwise be devoted to covering Rubio, which in turn leads to lower support in polls than Rubio would have otherwise received.

[4] All models are estimated using the MSBVAR Package in `R` (Brandt and Davis 2015).

BSVAR requires that priors be set indicating beliefs about the behavior of the series. Previous work using BSVAR is almost entirely used to analyze yearly or monthly data (Brandt and Freeman 2006, 2009; Brandt, Colaresi and Freeman 2008; Sims and Zha 1998), so although we use the priors from these works as a starting point, we make some important modifications. First we set the λ_0 to relatively high value of 0.9 indicating that we expect the sample variance to be good indicator of the real variance. This is because we believe the daily daily data do an adequate job of capturing the underlying variance. Another major change is setting λ_3 to 0.25. We do this as we believe it is feasible for the effects of past days to be relatively important in predicting the current day and that the effect will decrease relatively slowly. μ_5 and μ_6 are also of interest in this analysis as they reflect expectations about co-integration and common trends. Because we think that co-integration and common trends to be very possible we set these both to 5.

The values of the remaining hyperpriors and our reasoning for those values can be found in Table 2. In Section 3 we explore the role of these priors by setting them to other plausible values and re-estimating the full model. We do not find any substantial changes in our inferences.

Table 2: Function and Specification of Hyperpriors

	Function	Value	Reasoning
λ_0	Error covariance matrix scale	0.9	We expect sample errors to be fairly representative of true error.
λ_1	Persistence of series temporal dynamics	0.5	There will be relatively high persistence across days.
λ_3	Rate of lag variance decay	0.25	Given daily measures we expect that there will be a slow decay in the effect of previous days.
λ_4	Intercept standard deviation	0.25	We leave it possible for there to be a long-run average level in a series.
λ_5	Exogenous variable standard deviation	0.20	We expect the debates to have some effect.
μ_5	Autoregressive coefficient sum component	5	We expect it is possible that the series are difference stationary.
μ_6	Correlation of coefficients component	5	There is potential for a common trend.

Table 2: This table, based on Table 2 of Brandt and Freeman (2006, pg. 120), shows the specification of priors for our Bayesian Structural VARs. See Brandt and Freeman (2006, 2009) for more detail about how each prior affects the model.

Estimates are derived from draws from the posterior using a Gibbs sampler.[5] Convergence was verified through visual inspection of traceplots, autocorrelation functions, and the Geweke diagnostic (Geweke 1992).[6] These diagnostics were consistent with convergence.

[5]Gibbs sampling allows us explore the posterior through a process of iterated sampling of a subset of parameters, conditional on the remaining parameters and the data (Jackman 2009, pg. 214-215)

[6]We used the Gibbs sampler in MSBVAR to run 1 chain of 44,000 draws, and discarded the first 4,000 as

In addition to the endogenous variables, we included the primary debates as exogenous variables. Each debate is represented by a binary variable equal to 1 for the day of the debate and 0 for all other days. There were a total of 7 debates in our time frame. The debate variables capture sudden changes caused by the debates. Including exogenous debate indicators is necessary because political debates can cause sudden fluctuations in public interest which could otherwise be misattributed to media coverage (or vice versa).

2.3 Lag Lengths

As discussed, we estimate each model individually. After estimating each model we tested to see if the residuals were consistent with a white noise pattern. When the residuals were not white noise we adjusted the number of lags in the model. We found that 7 lags led to white noise residuals in almost all of the models except for in the case of Jeb Bush and Ben Carson. In these two cases we needed to include 10 lags. This is detailed in Table 3 which displays the number of lags in each model presented in the text.

Table 3: Number of Lags per

Candidate	Lags
Trump	7
Bush	10
Carson	10
Christie	7
Cruz	7
Fiorina	7
Paul	7
Rubio	7

Table 3: Number of lags in each of the model presented in the main text.

2.4 Residuals

We used Ljung-Box test for white noise residuals (Ljung and Box 1978). In Table 4 we display p-values for these tests with 14 lags. The lack of significant values indicate that residuals were white noise processes.

burn-in. The likelihood was then normalized using the method recommended by Waggoner and Zha (2003).

Table 4: Ljung-Box p-values for Residuals

Candidate	Media	Interest	Support
Trump	0.79	0.63	0.92
Bush	0.27	0.23	0.99
Carson	0.97	0.98	0.41
Christie	0.79	0.37	0.29
Cruz	0.98	0.68	0.12
Fiorina	0.99	0.98	0.80
Paul	0.37	0.95	0.79
Rubio	0.86	0.93	0.40

Table 4: P-values from Ljung-Box tests with 14 lags.

3 Robustness to Priors

To examine the effects of the priors we picked a few potential alternative values and re-estimated the models to examine if this would lead to any substantial changes in our results. In the subsections below we show the original IRFs alongside the IRFs for several of these robustness checks. We did not find any substantial changes. The priors we changed were the ones that we considered to be potentially the most critical for our results.

3.1 Setting $\lambda_3 = .5$

Figure 1: Media → Support Robustness

Original

Robust to Priors

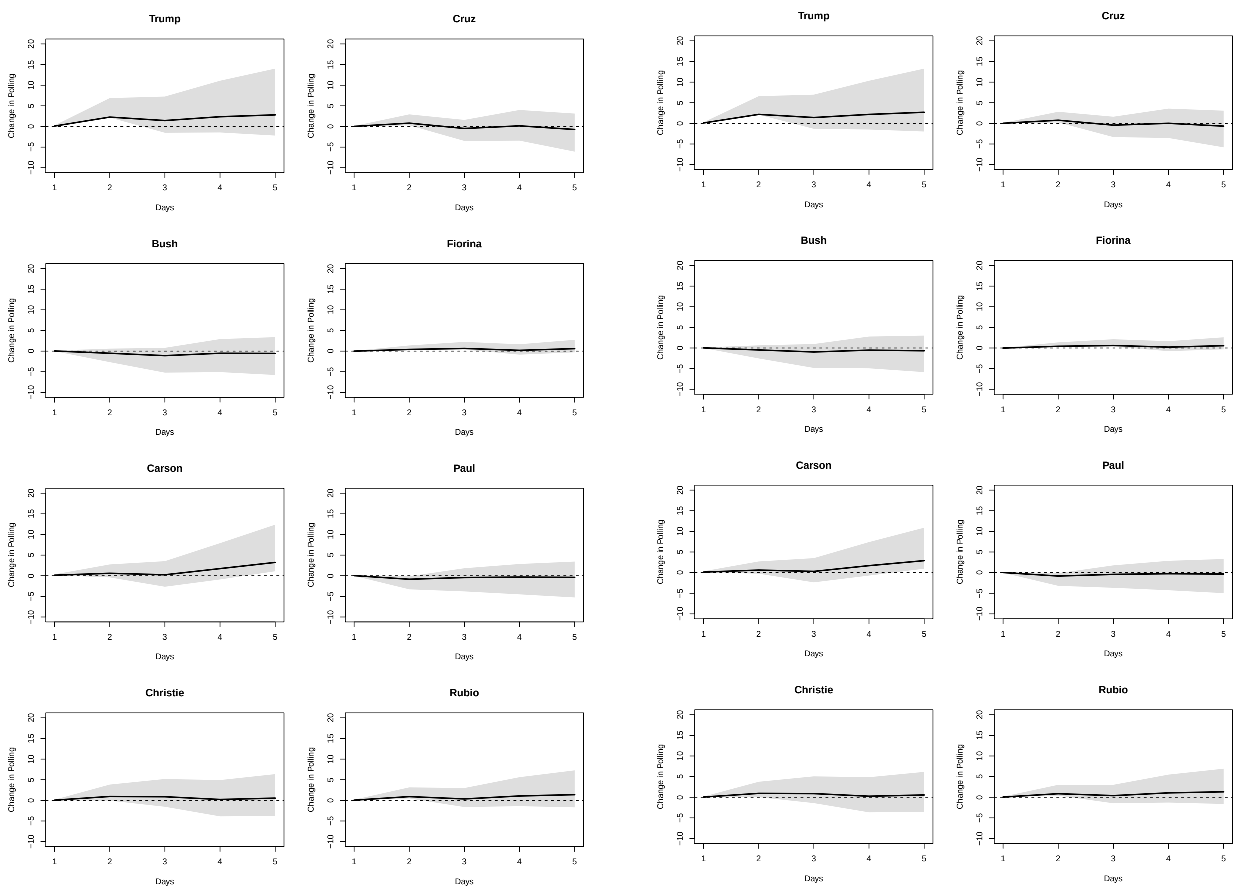

Figure 1: Comparing the IRFs of a shock to Media on Support presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 2: Support → Media Robustness

Original

Robust to Priors

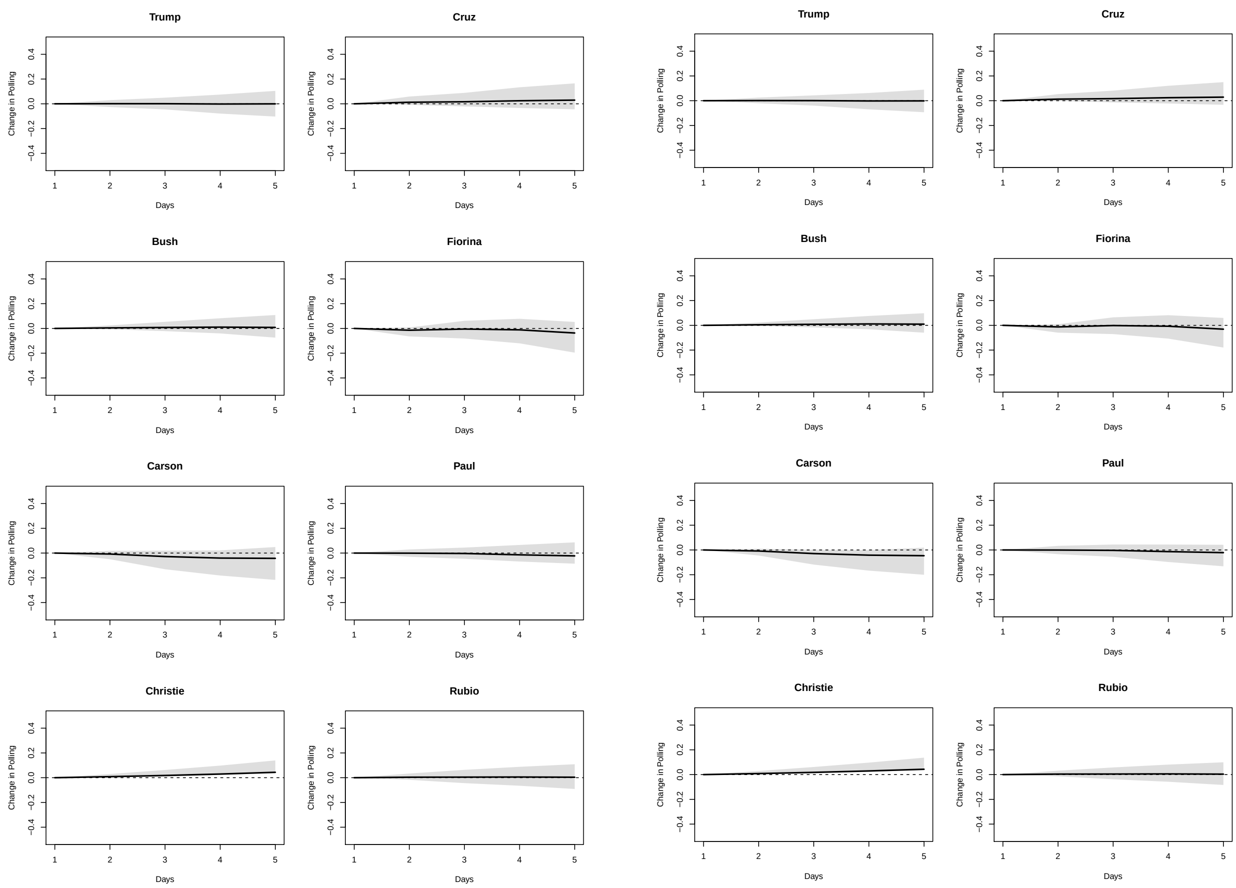

Figure 2: Comparing the IRFs of a shock to Support on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 3: Interest → Media Robustness

Original

Robust to Priors

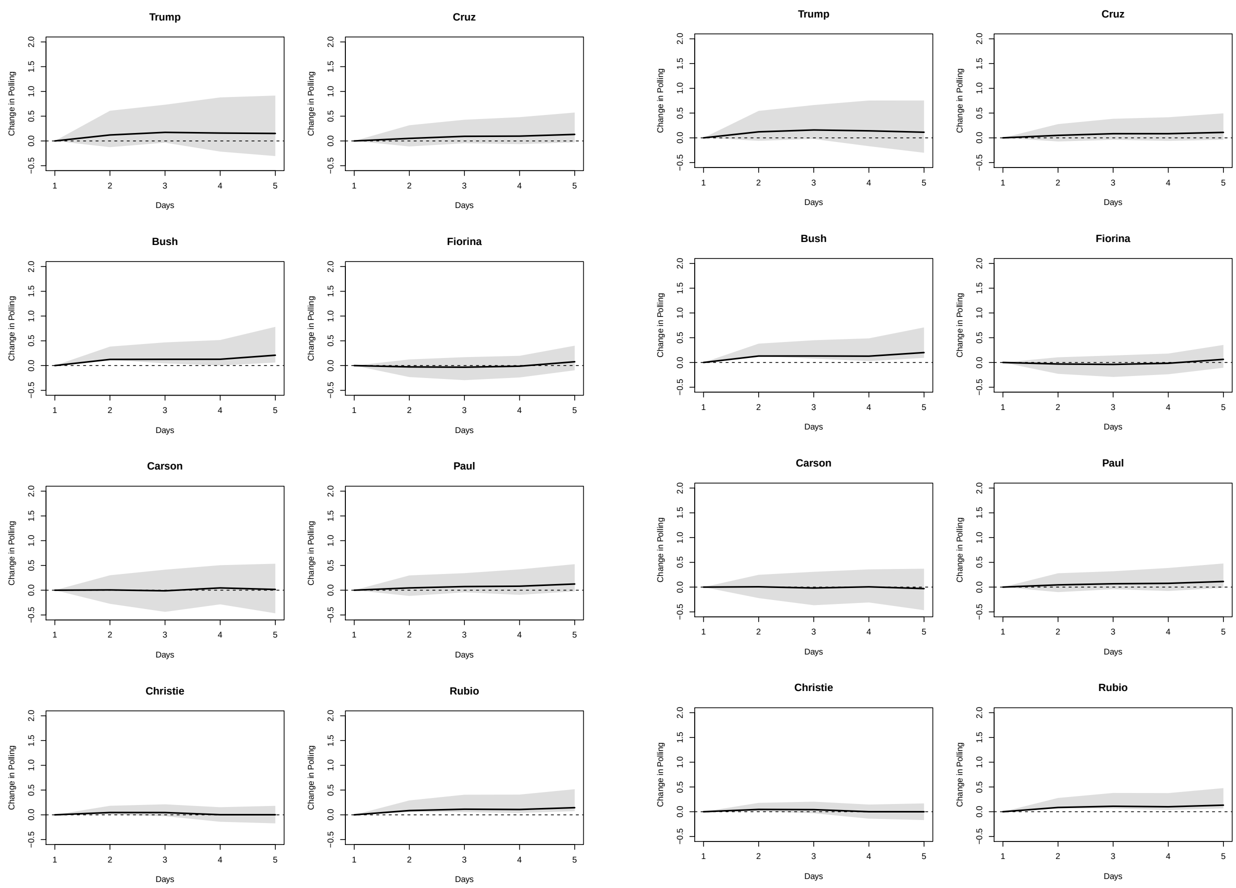

Figure 3: Comparing the IRFs of a shock to Interest on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 4: Interest → Support Robustness

Original

Robust to Priors

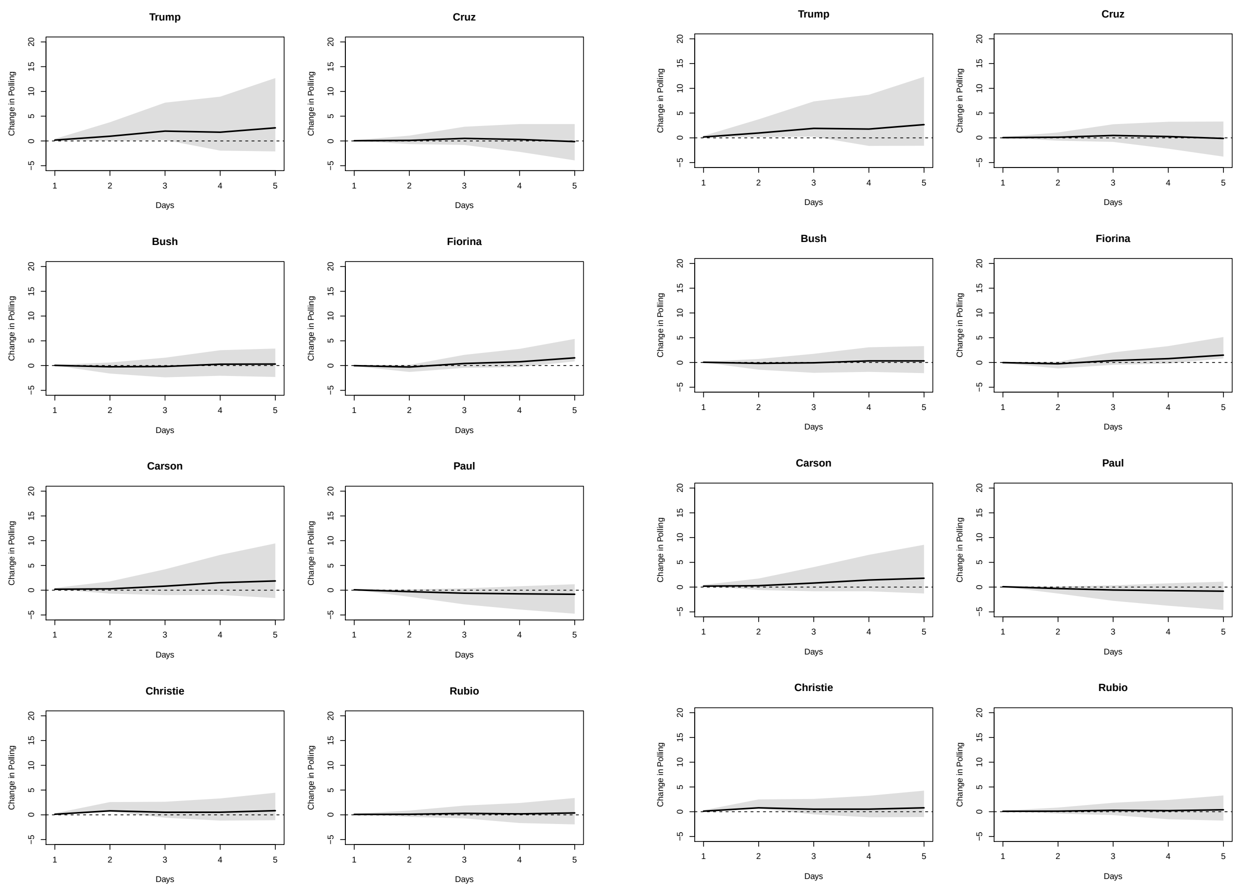

Figure 4: Comparing the IRFs of a shock to Interest on Support presented in the main text (left side) to IRFs with an alternative prior specification.

3.2 Setting $\lambda_3 = .1$

Figure 5: Media $\rightarrow$ Support Robustness

Original

Robust to Priors

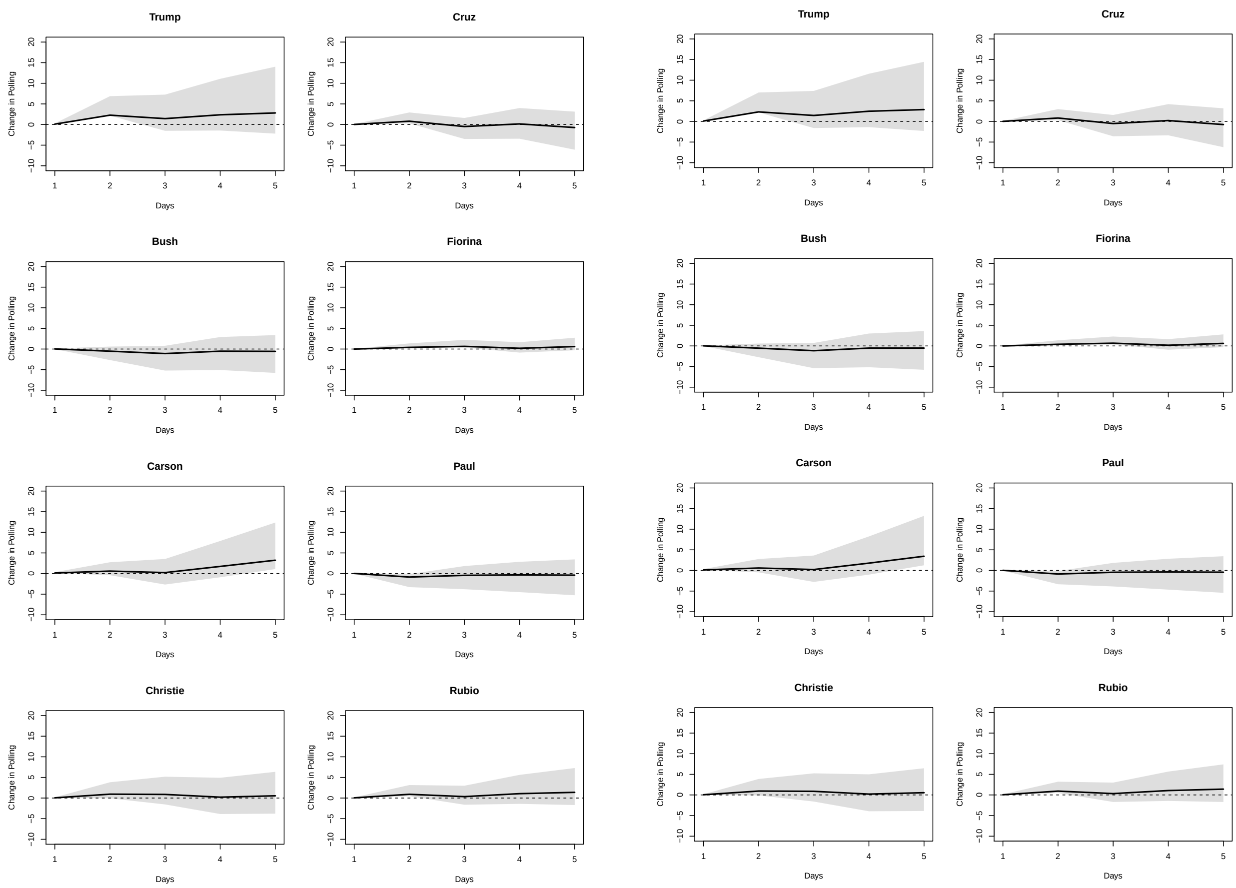

Figure 5: Comparing the IRFs of a shock to Media on Support presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 6: Support → Media Robustness

Original

Robust to Priors

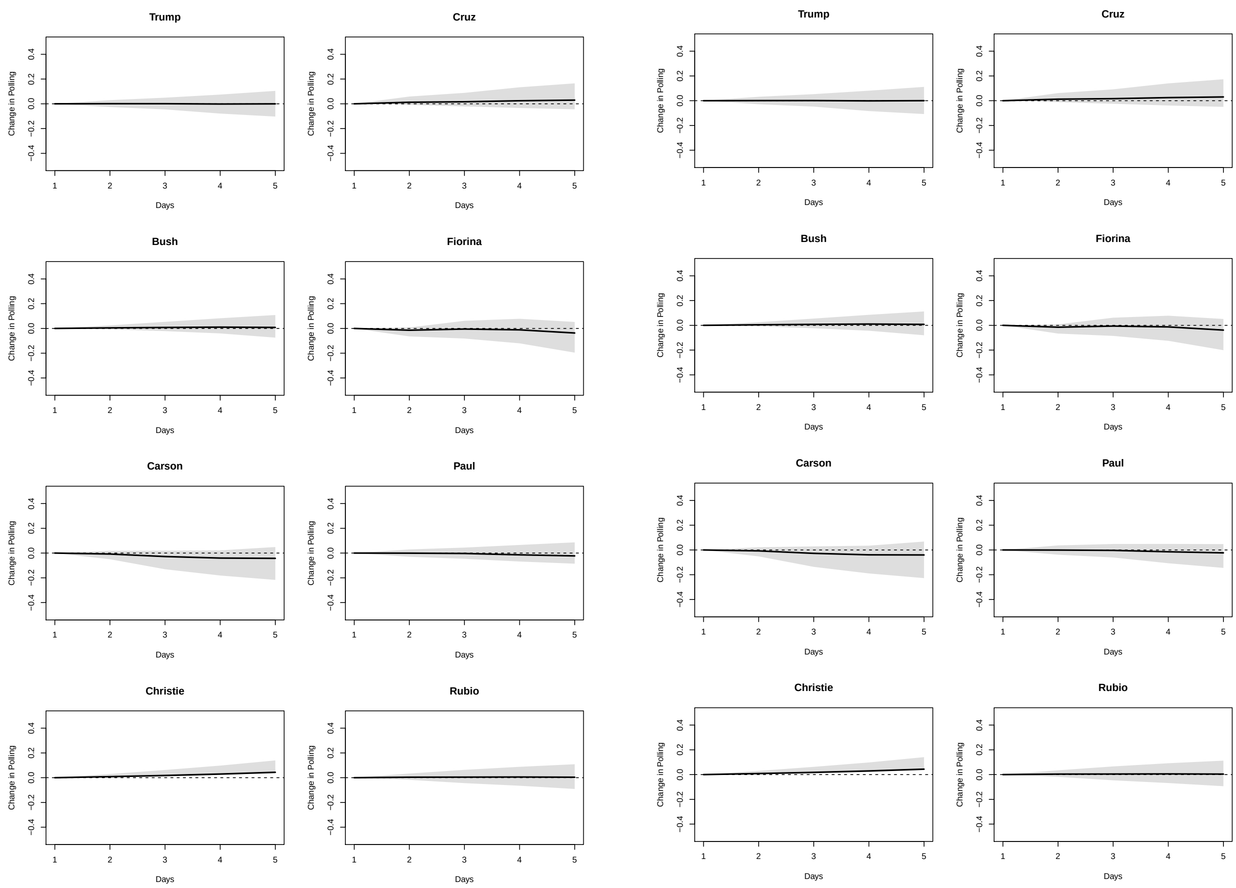

Figure 6: Comparing the IRFs of a shock to Support on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 7: Interest → Media Robustness

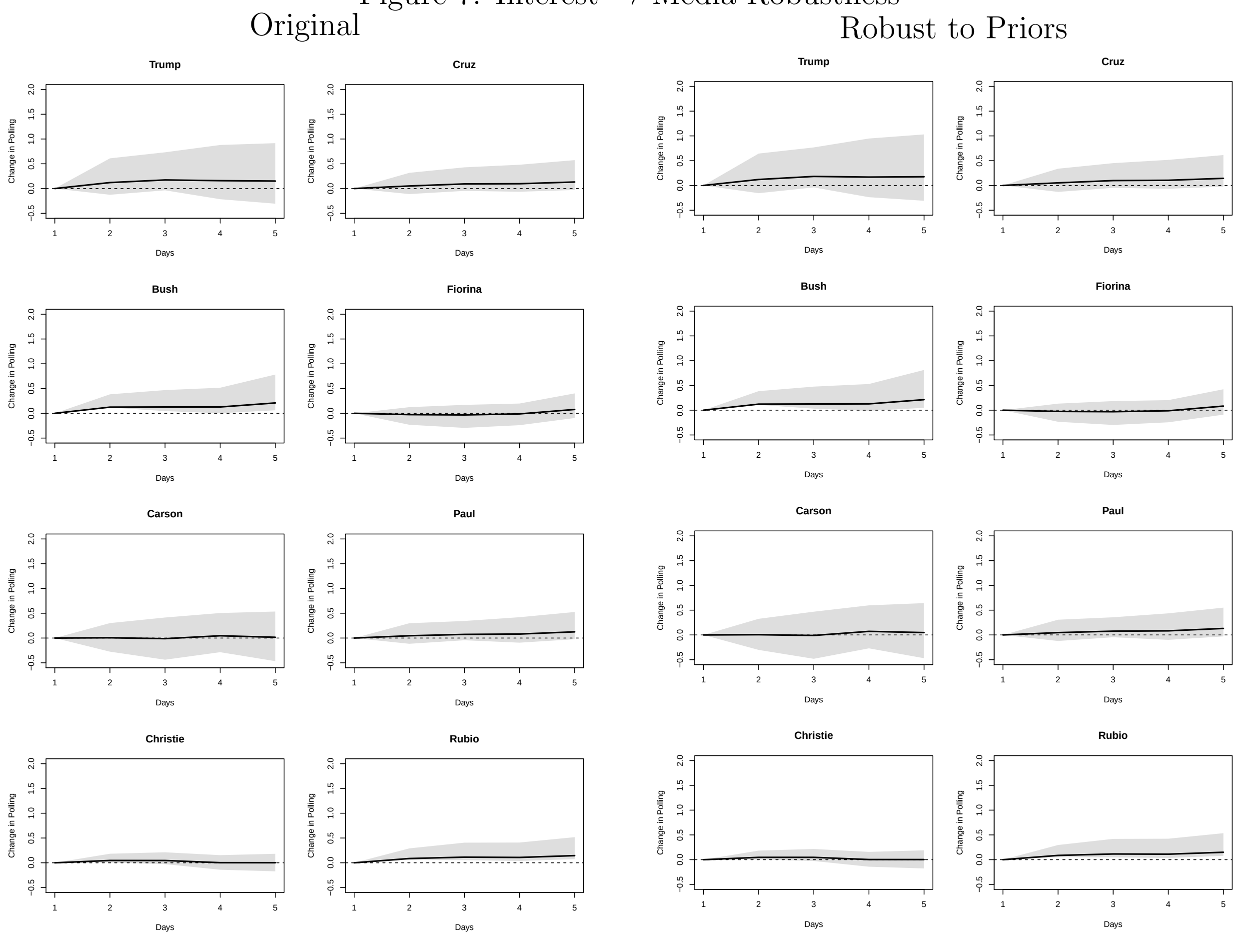

Figure 7: Comparing the IRFs of a shock to Interest on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 8: Interest → Support Robustness

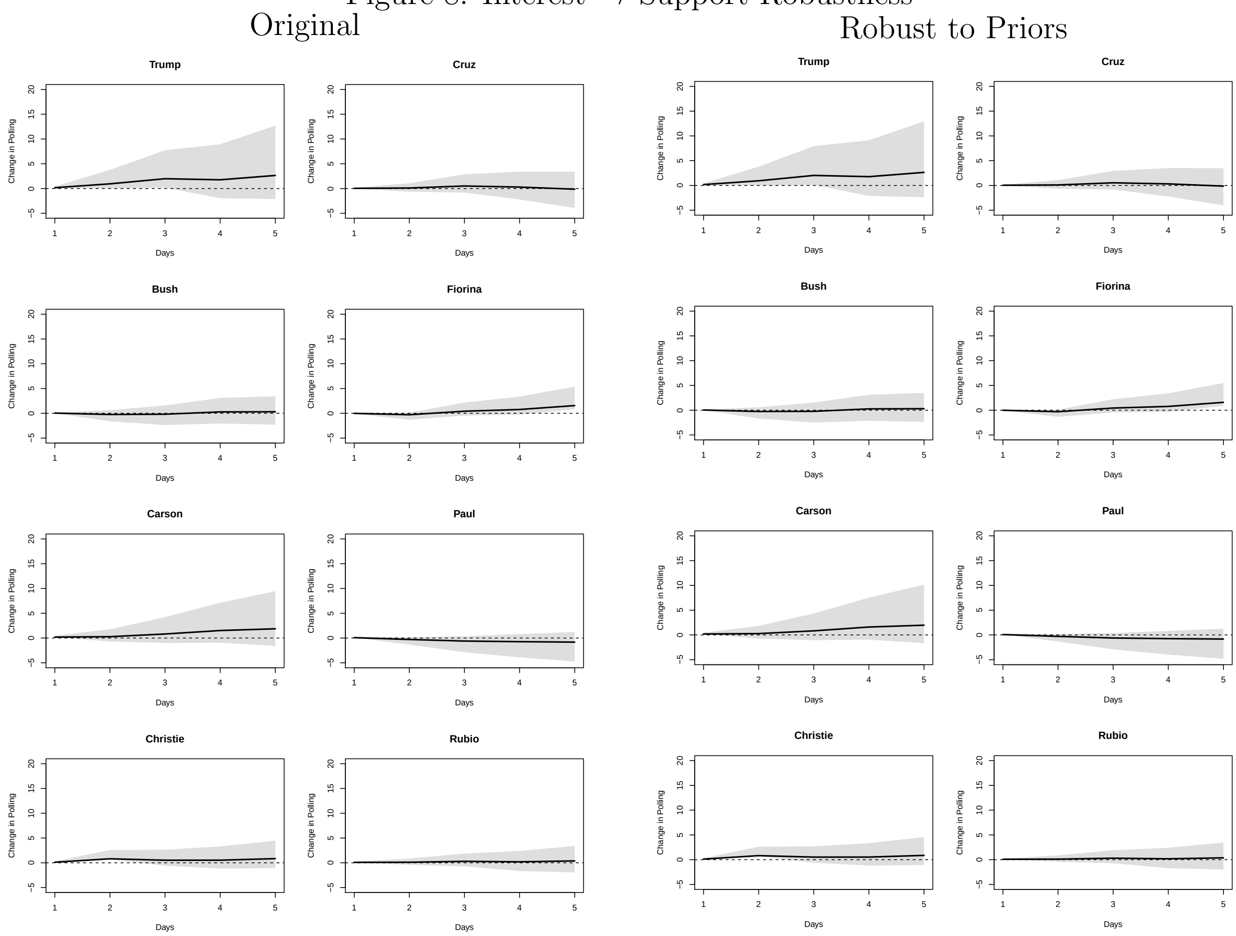

Figure 8: Comparing the IRFs of a shock to Interest on Support presented in the main text (left side) to IRFs with an alternative prior specification.

3.3 Setting $\mu_5 = \mu_6 = 1$

Figure 9: Media → Support Robustness

Original

Robust to Priors

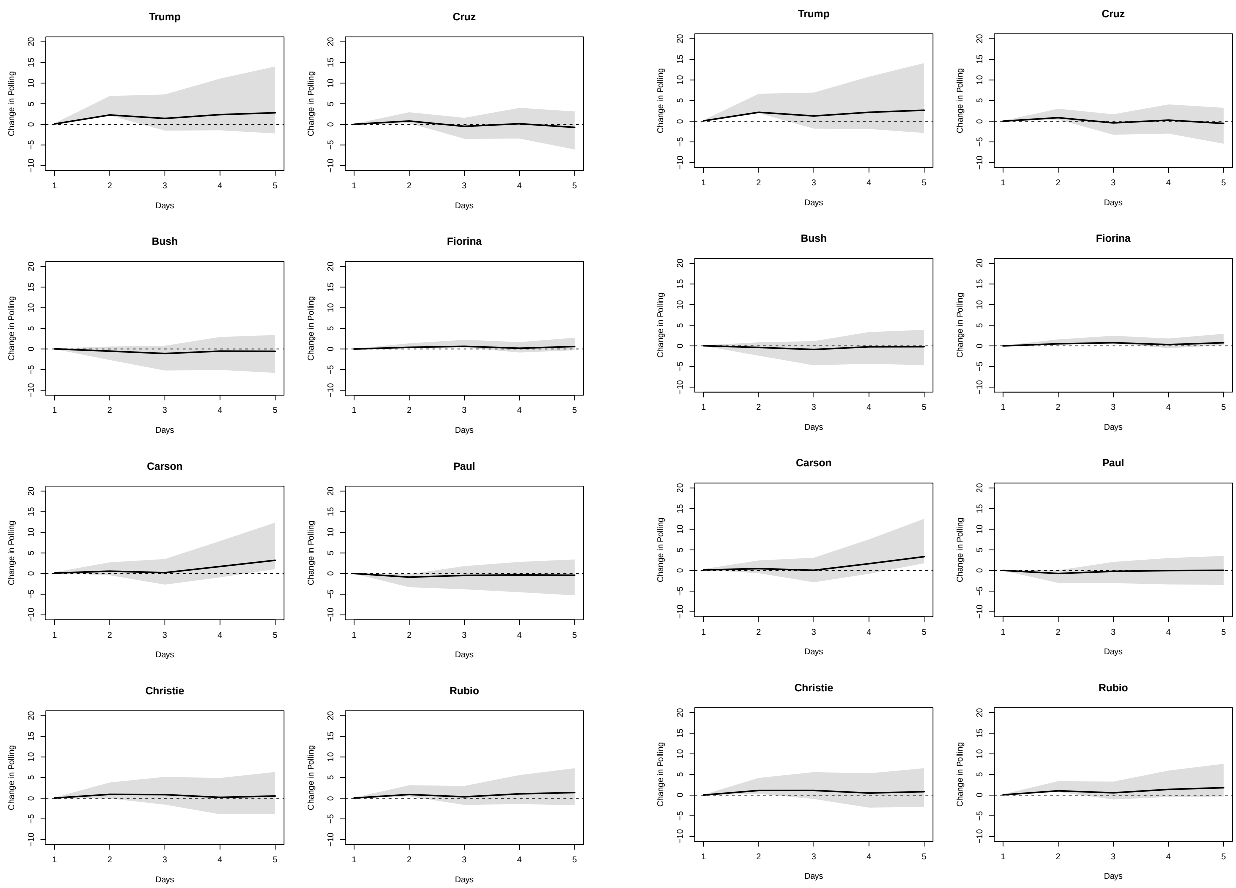

Figure 9: Comparing the IRFs of a shock to Media on Support presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 10: Support → Media Robustness

Original

Robust to Priors

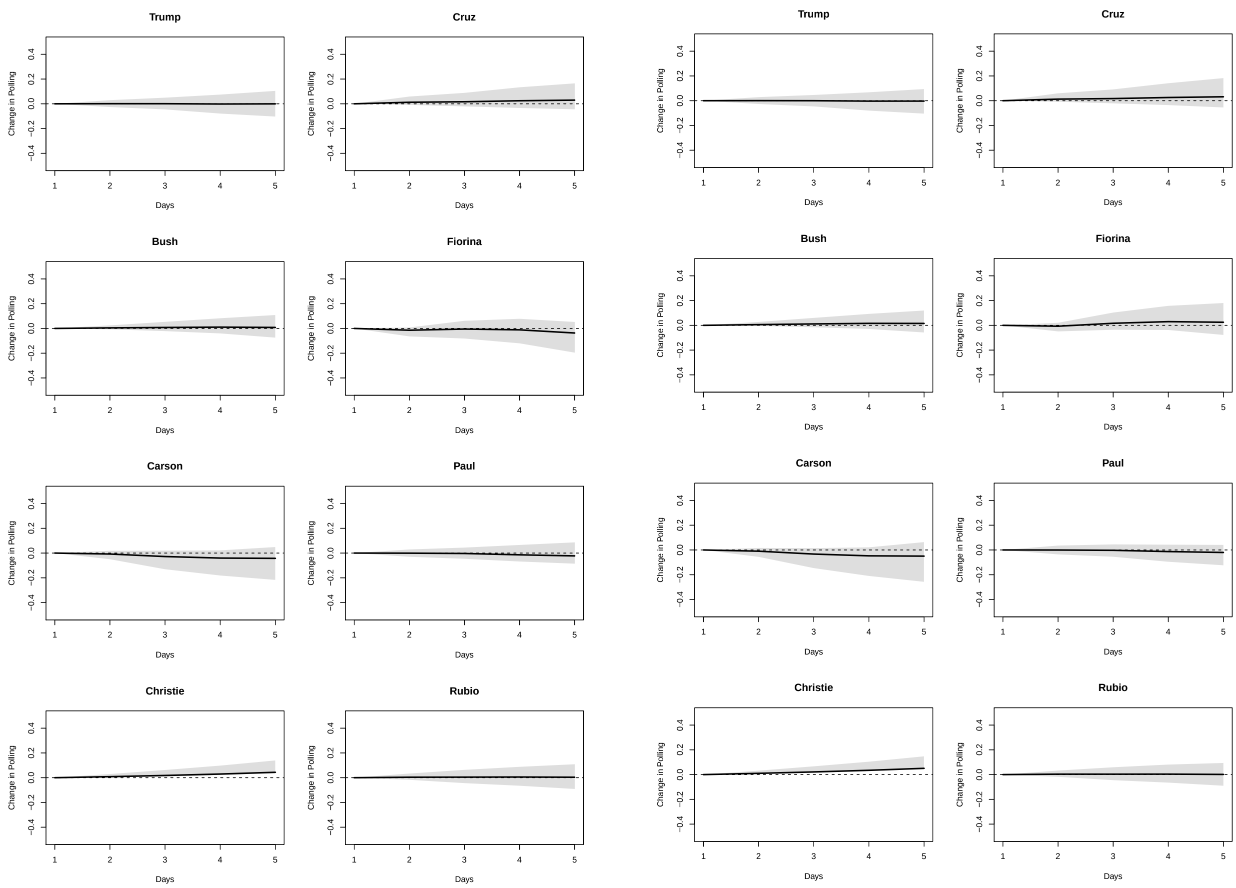

Figure 10: Comparing the IRFs of a shock to Support on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 11: Interest → Media Robustness

Original

Robust to Priors

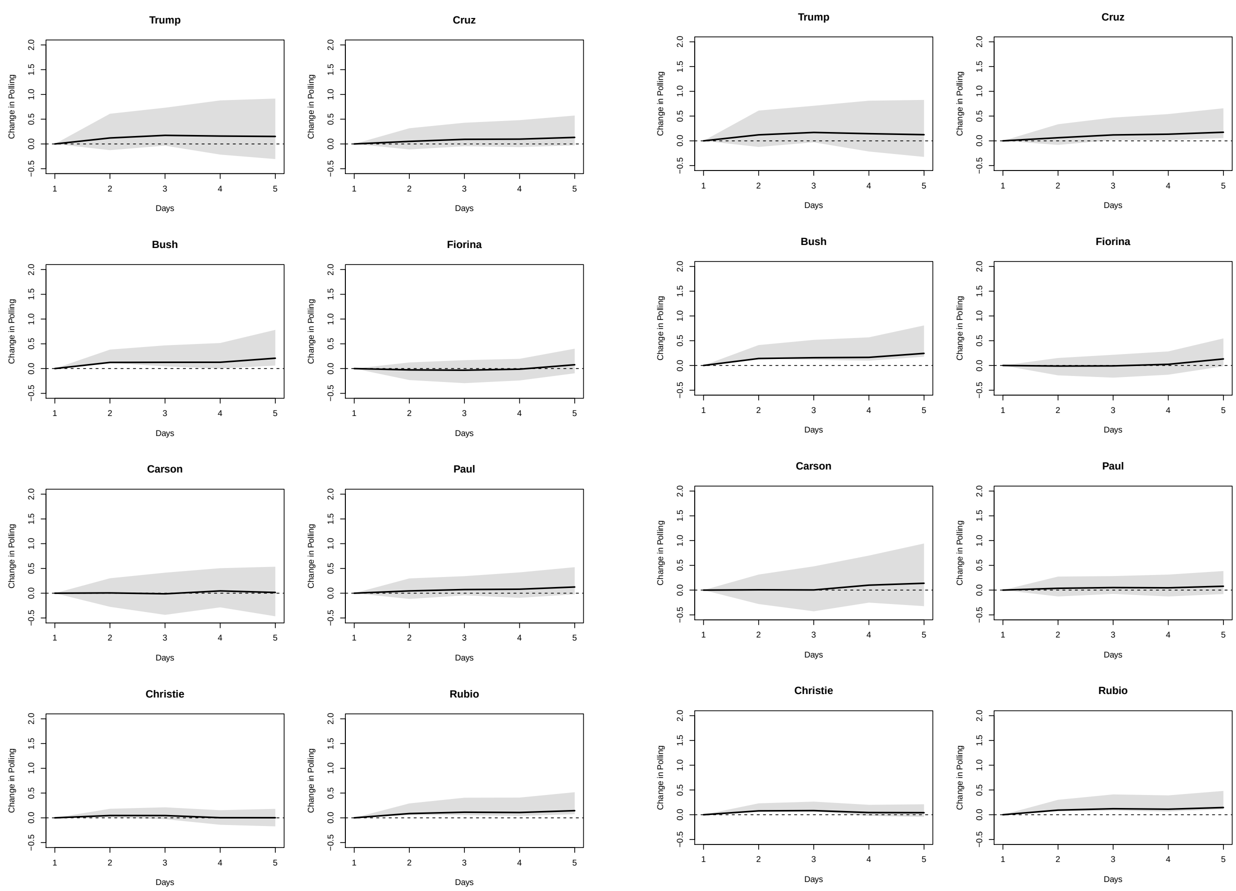

Figure 11: Comparing the IRFs of a shock to Interest on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 12: Interest → Support Robustness

Original

Robust to Priors

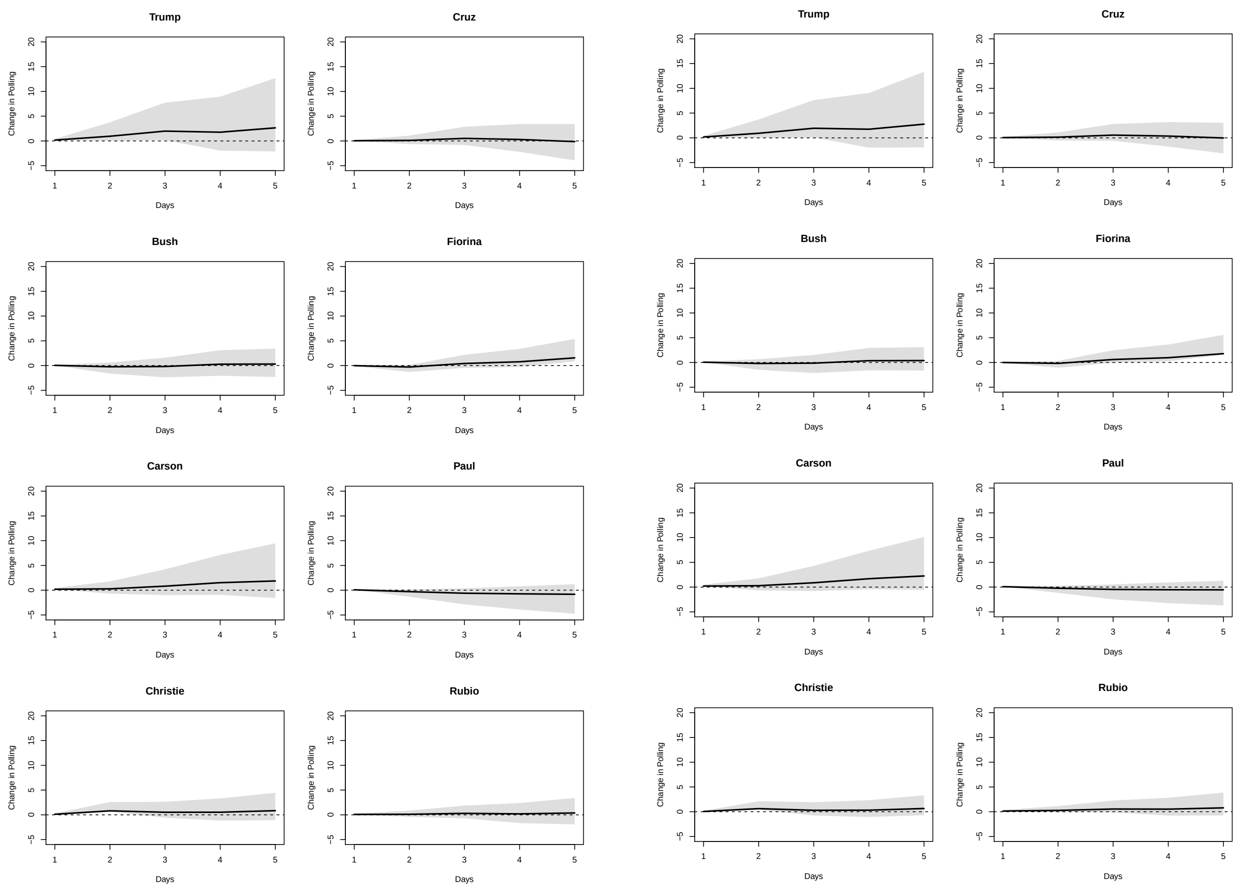

Figure 12: Comparing the IRFs of a shock to Interest on Support presented in the main text (left side) to IRFs with an alternative prior specification.

3.4 Setting $\lambda_1 = .5$

Figure 13: Media → Support Robustness

Original

Robust to Priors

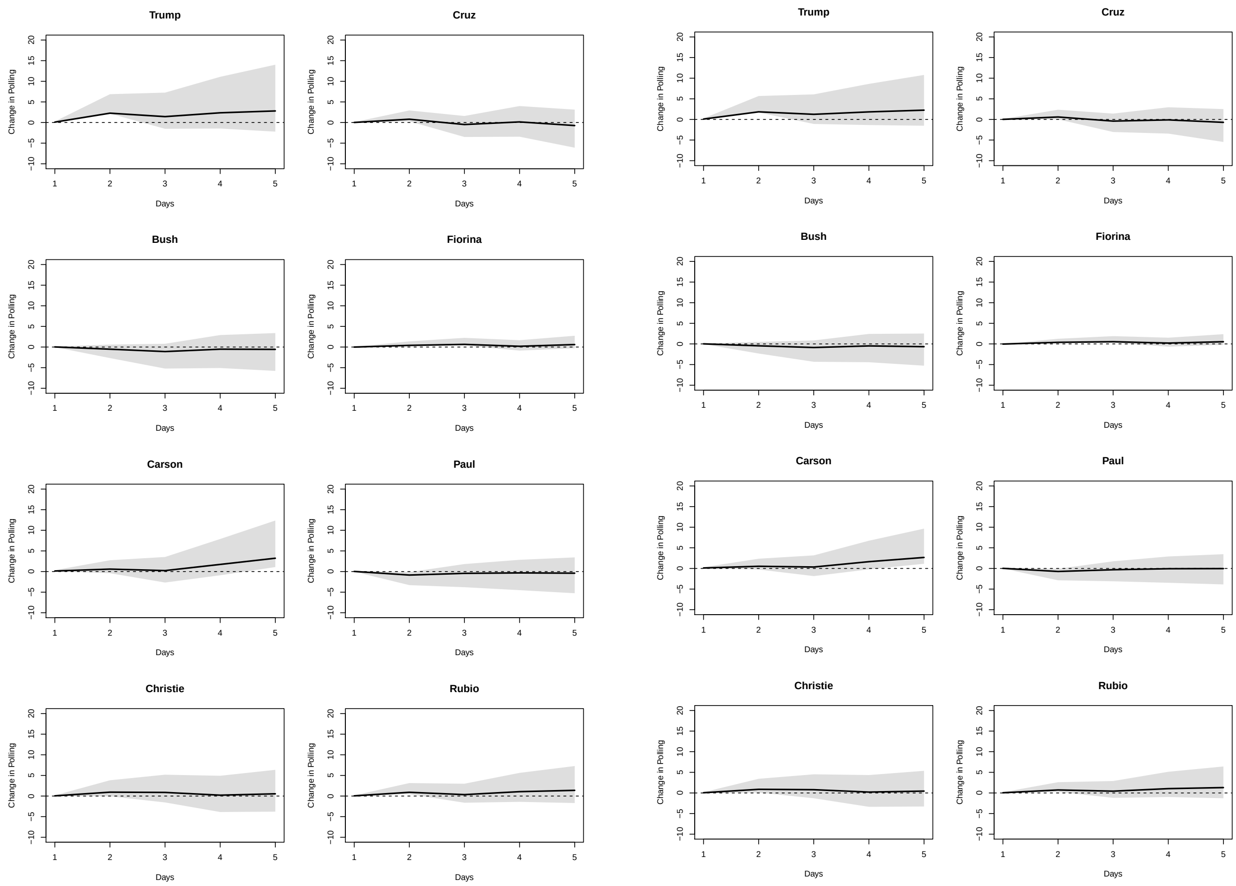

Figure 13: Comparing the IRFs of a shock to Media on Support presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 14: Support → Media Robustness

Original

Robust to Priors

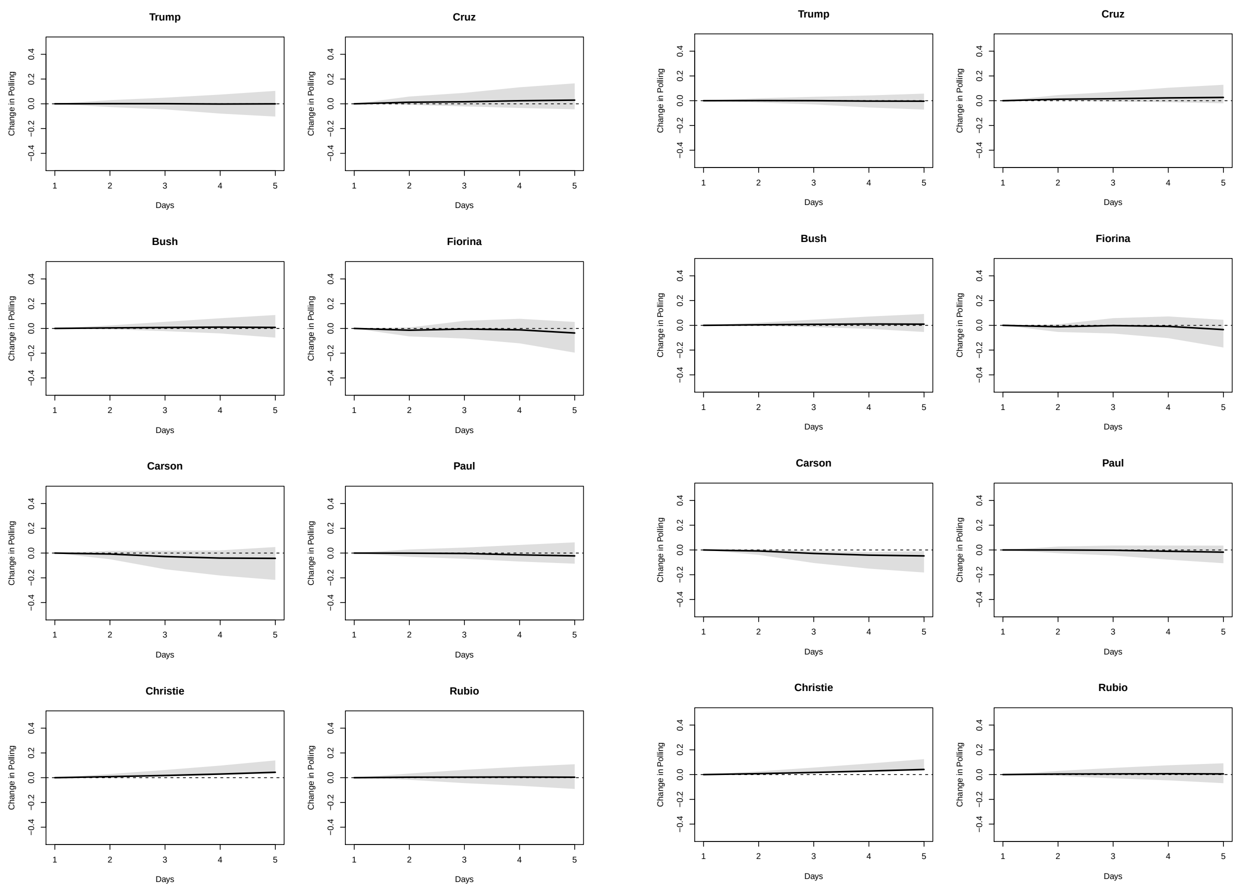

Figure 14: Comparing the IRFs of a shock to Support on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 15: Interest → Media Robustness

Original

Robust to Priors

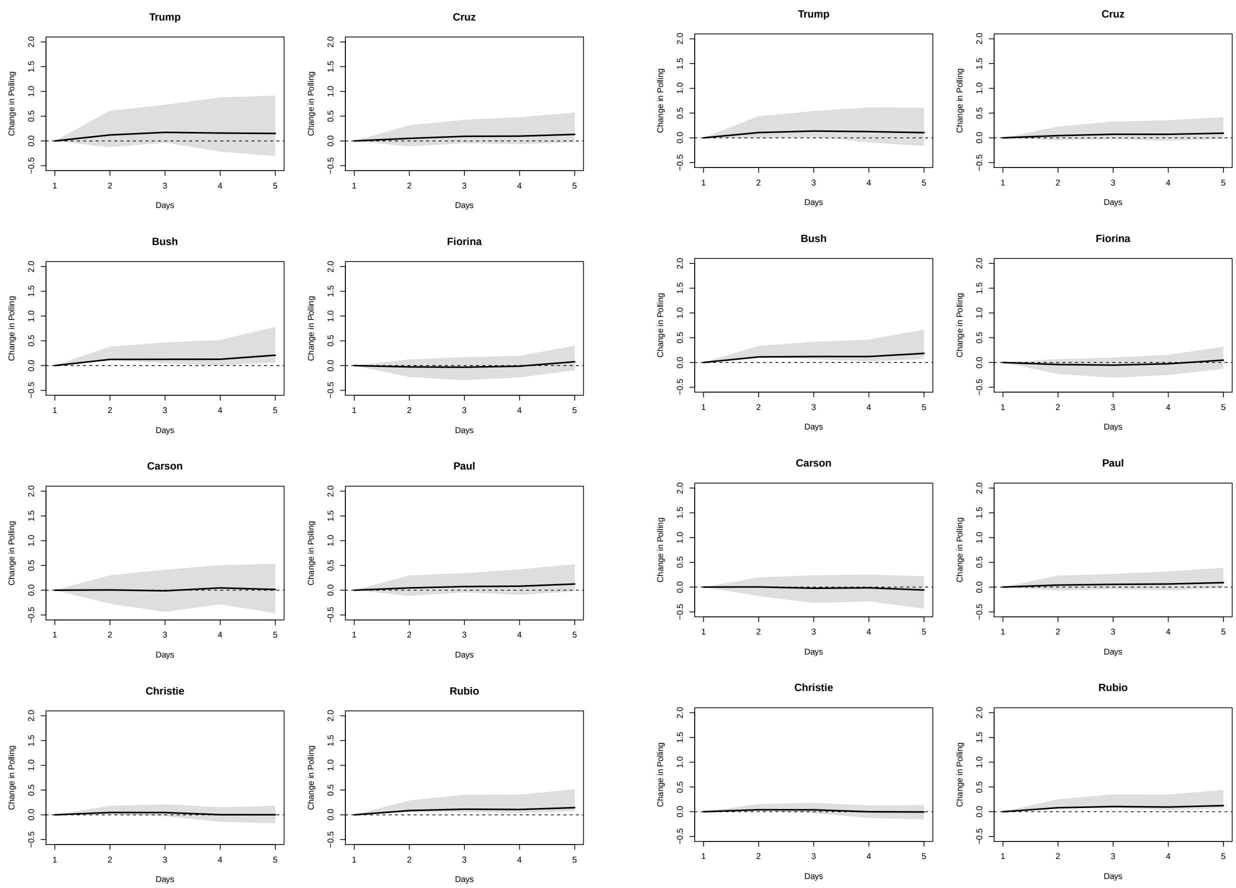

Figure 15: Comparing the IRFs of a shock to Interest on Media presented in the main text (left side) to IRFs with an alternative prior specification.

Figure 16: Interest → Support Robustness

Original

Robust to Priors

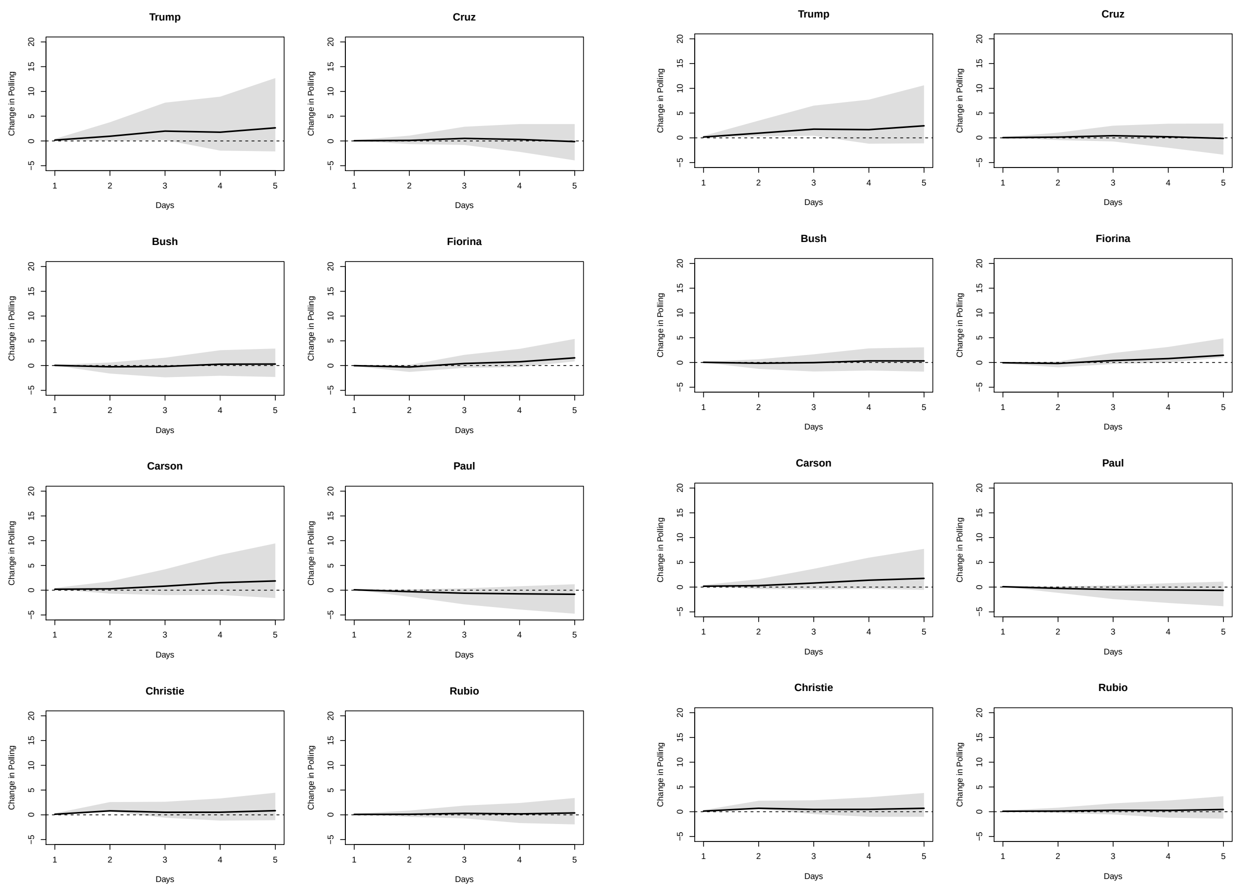

Figure 16: Comparing the IRFs of a shock to Interest on Support presented in the main text (left side) to IRFs with an alternative prior specification.

Table 5: Alternative Contemporaneous Identification

		Media	Interest	Polling
Media	=	X		
Interest	=	X	X	
Polling	=	X		X

Table 5: This table displays the identification of the contemporaneous shocks. Each row represents an equation with the label being the dependent variable. The columns indicate the contemporaneous variable. An X means that the coefficient was estimated, while a blank means that the coefficient was constrained to 0. The diagonals are never constrained to zero by definition because they are used to include the dependent variable. See text for details

4 Robustness to Identification

In this section we imposed an alternative identification strategy were only media has immediate effects. This identification is drawn from public discourse around elections which often frames news media as the primary driver of election dynamics. We again recreate the IRFs alongside the new IRFs from this alternative model. We find that there are no substantiative changes.

Figure 17: Media → Support Robustness

Original

Robust to Identification

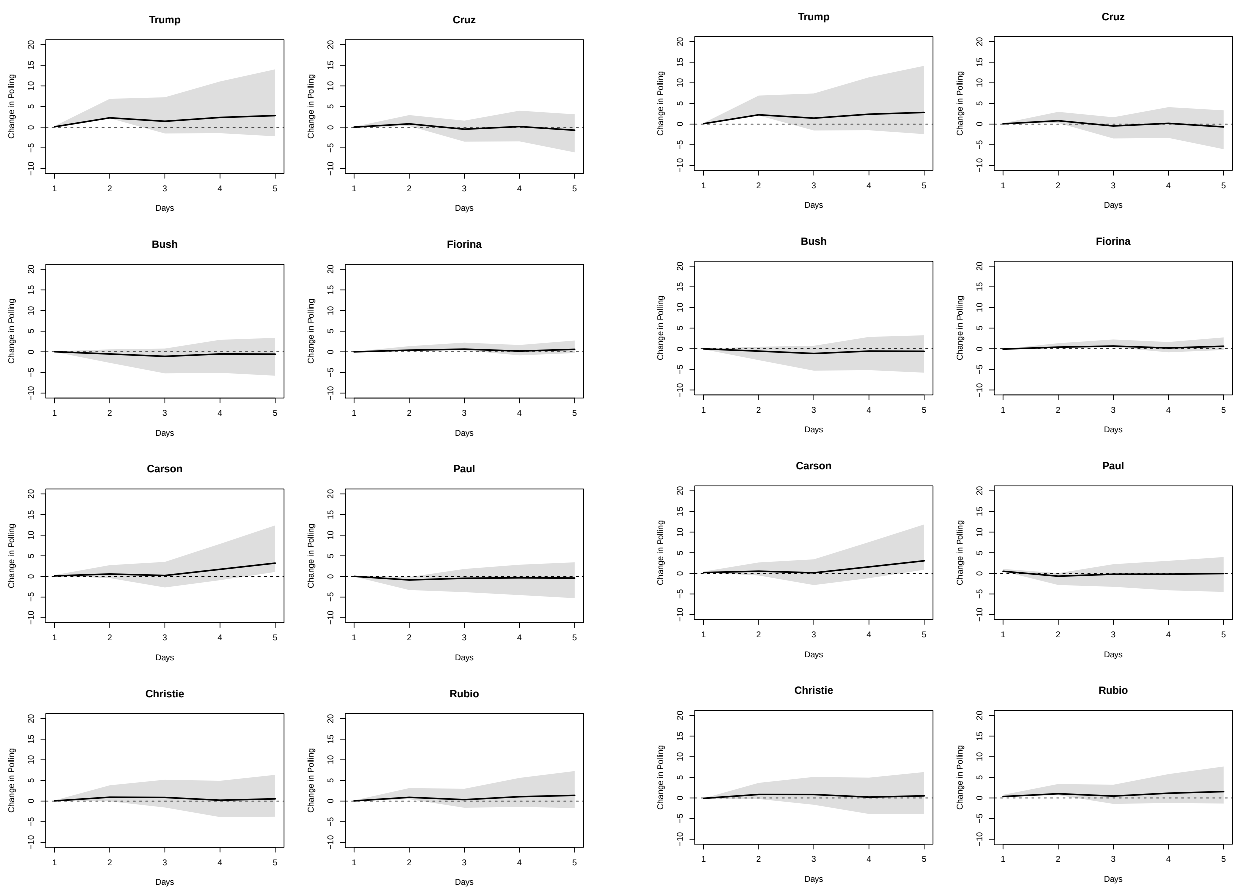

Figure 17: Comparing the IRFs of a shock to Media on Support presented in the main text (left side) to IRFs with an alternative identification specification.

Figure 18: Support → Media Robustness

Original

Robust to Identification

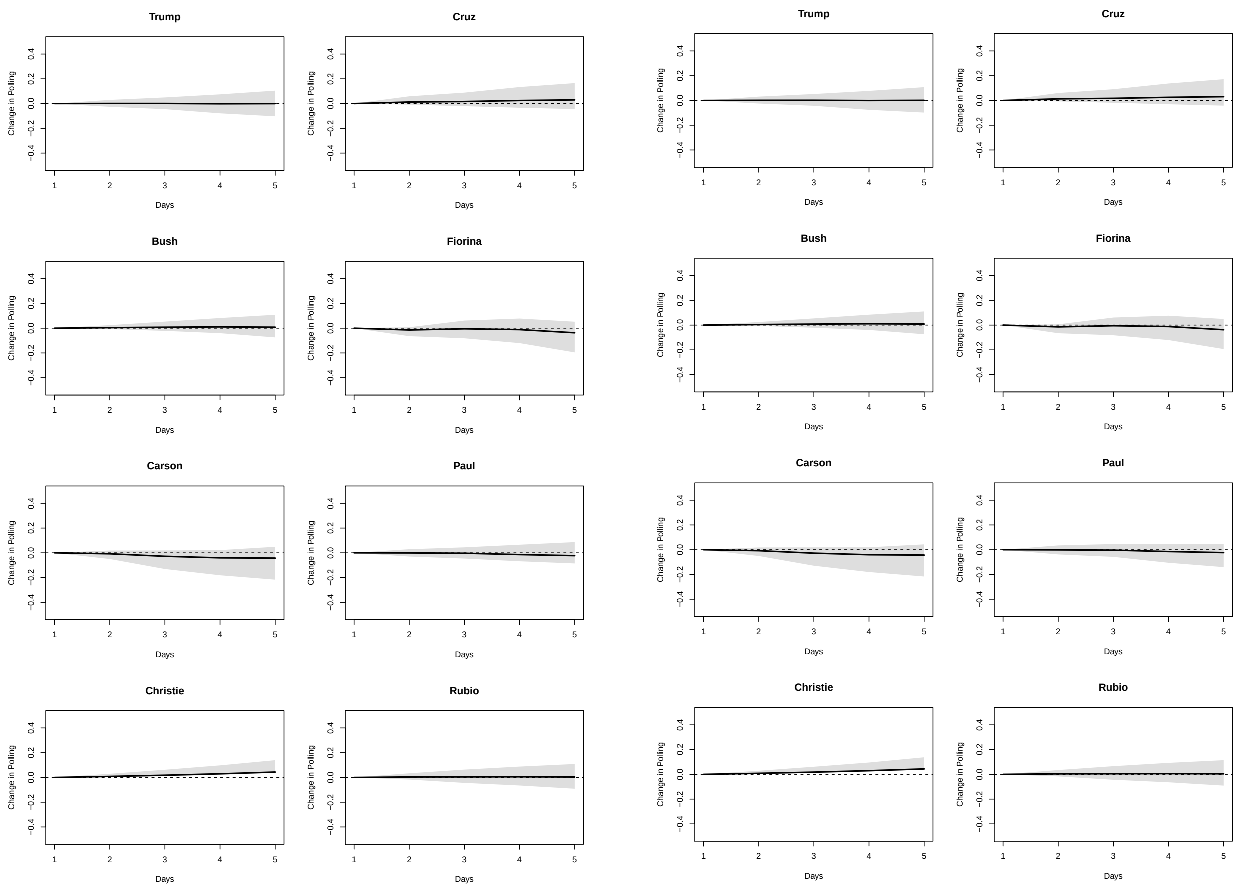

Figure 18: Comparing the IRFs of a shock to Support on Media presented in the main text (left side) to IRFs with an alternative identification specification.

Figure 19: Interest → Media Robustness

Original

Robust to Identification

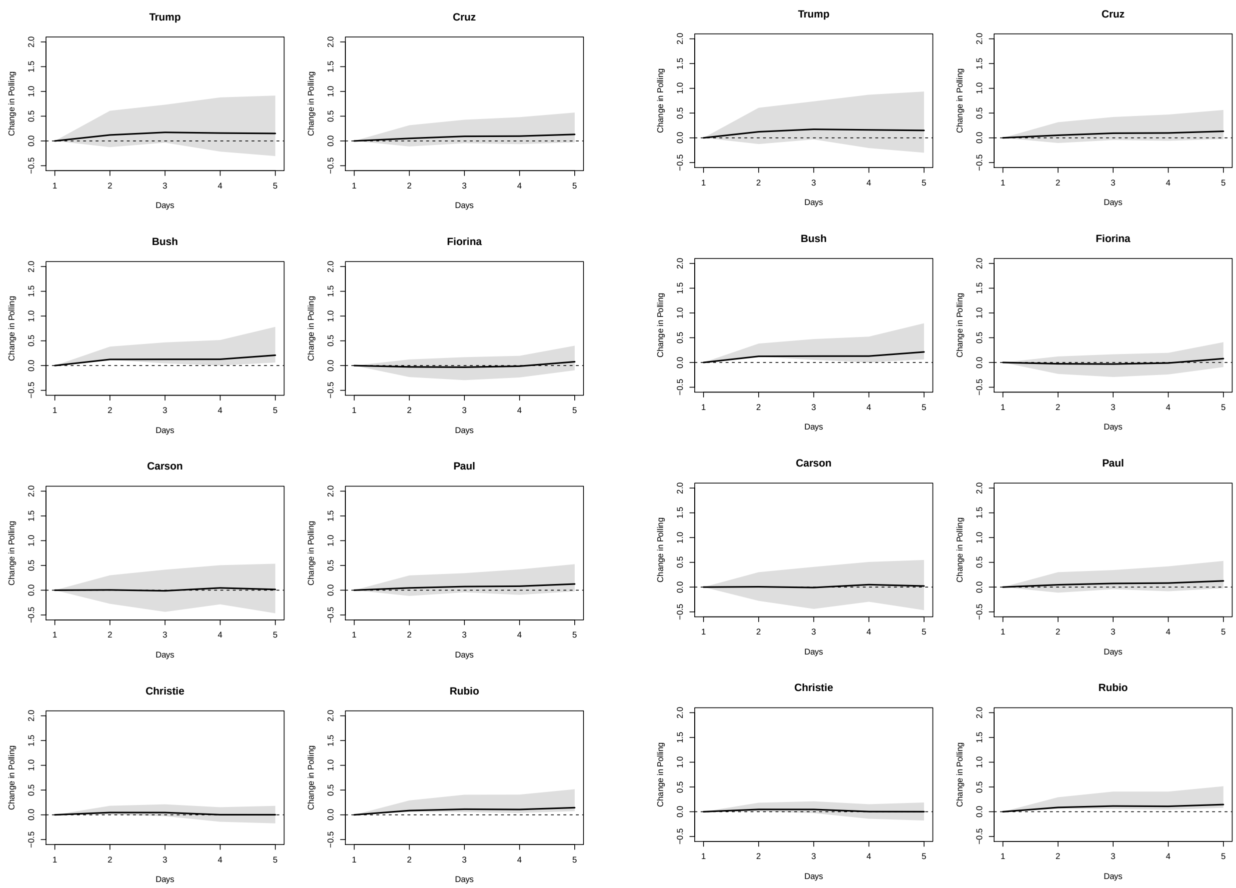

Figure 19: Comparing the IRFs of a shock to Interest on Media presented in the main text (left side) to IRFs with an alternative identification specification.

Figure 20: Interest → Support Robustness

Original

Robust to Identification

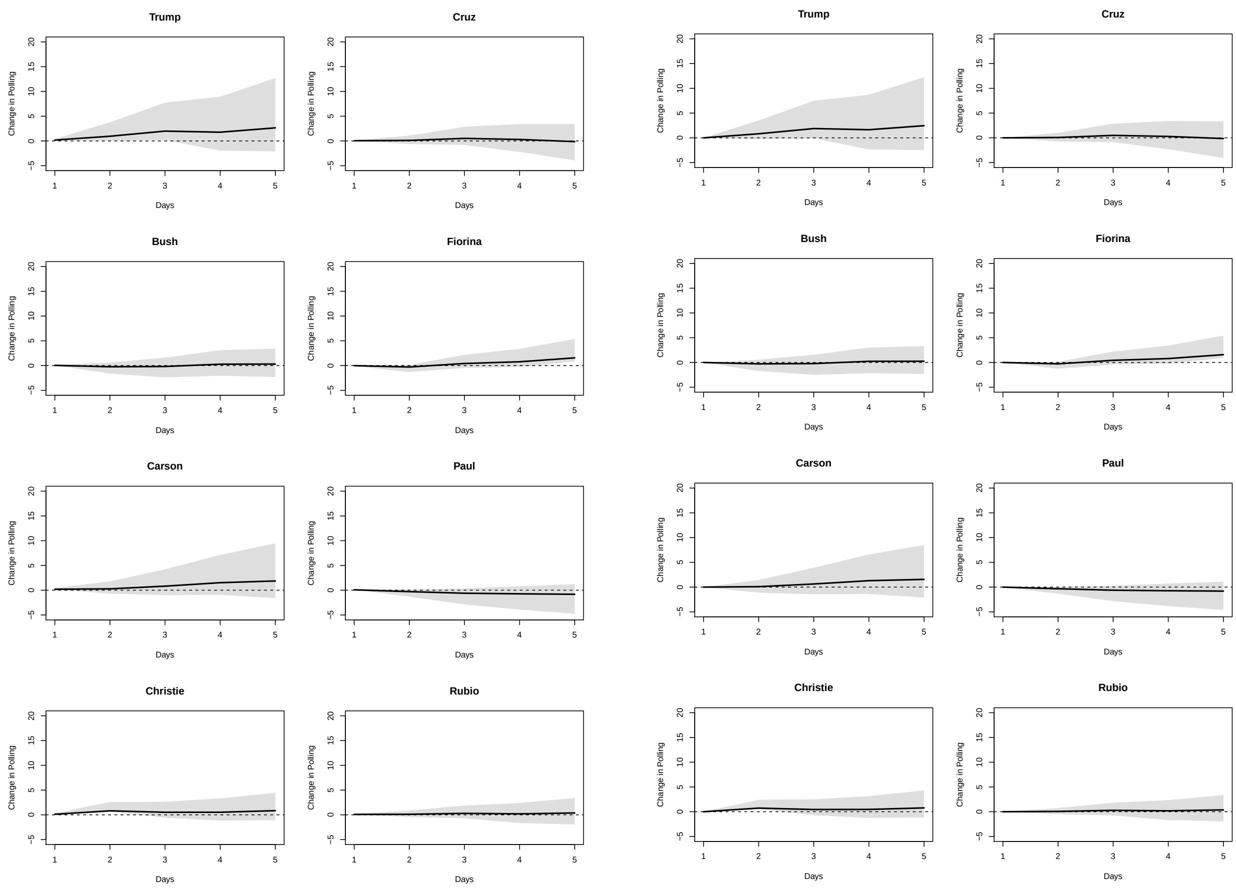

Figure 20: Comparing the IRFs of a shock to Interest on Support presented in the main text (left side) to IRFs with an alternative identification specification.

5 Robustness to Polling Aggregation

In this final robustness check we re-estimate the models using a different assumption to aggregate support for candidates. Instead of locating each poll in the middle of the timeframe for when it was fielded, we set it to the very first day it was in the field. This allows us to check to see if, by selecting the middle of the fielding range, we were artificially lagging the support scores. In these estimates we find that our credible intervals are wider but the inferences are similar to what was presented in the text.

Figure 21: Media → Support Robustness

Original | Robust to Polling Aggregation

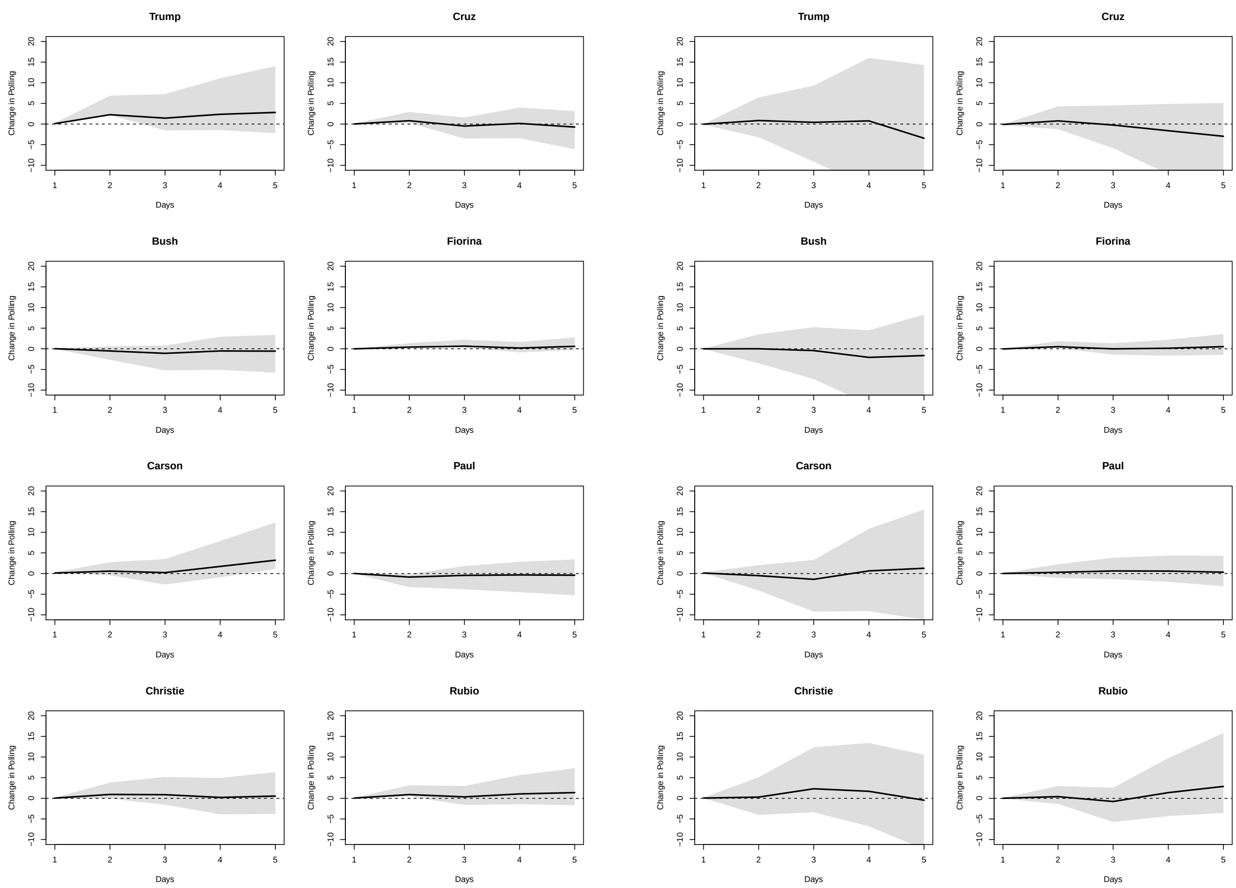

Figure 21: Comparing the IRFs of a shock to Media on Support presented in the main text (left side) to IRFs with an alternative identification specification.

Figure 22: Support → Media Robustness

Original | Robust to Polling Aggregation

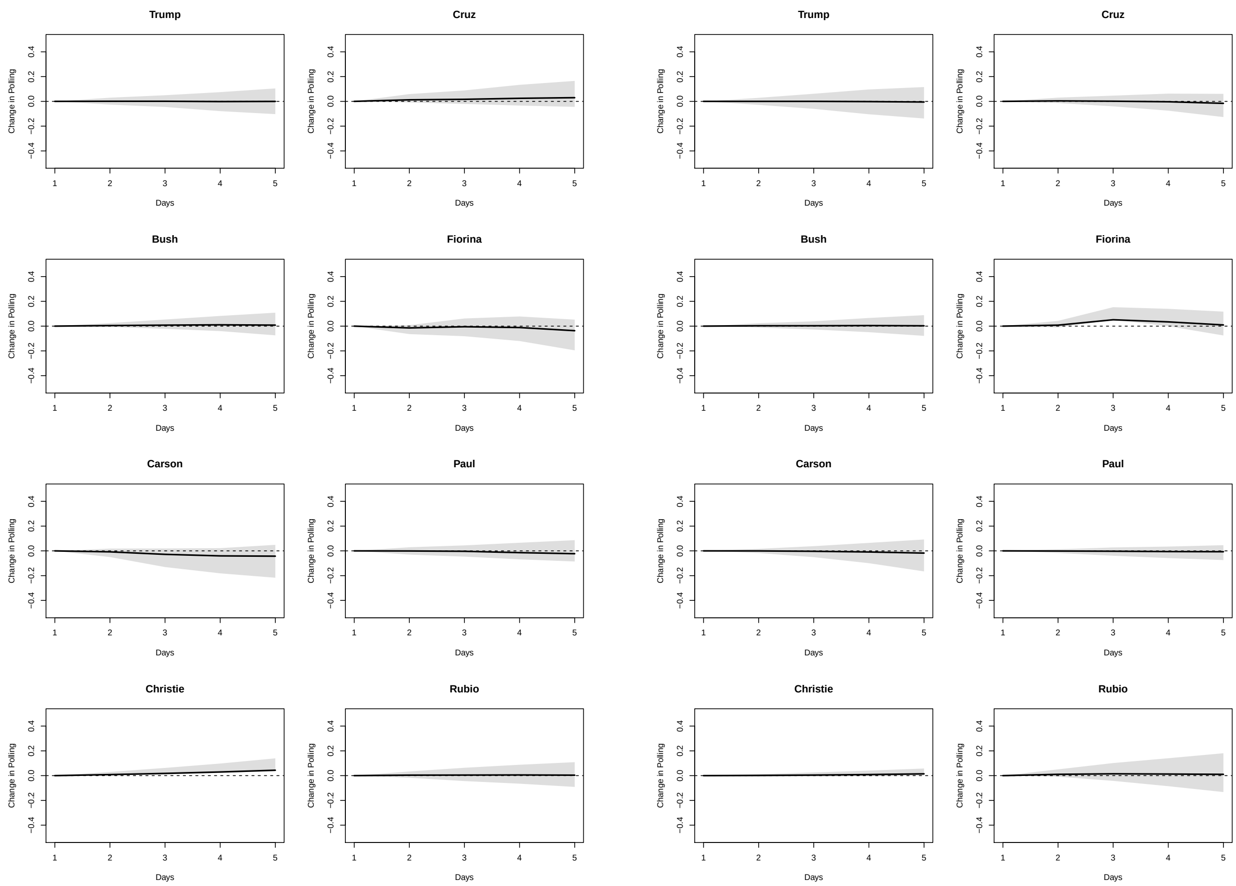

Figure 22: Comparing the IRFs of a shock to Support on Media presented in the main text (left side) to IRFs with an alternative identification specification.

Figure 23: Interest → Media Robustness

Original | Robust to Polling Aggregation

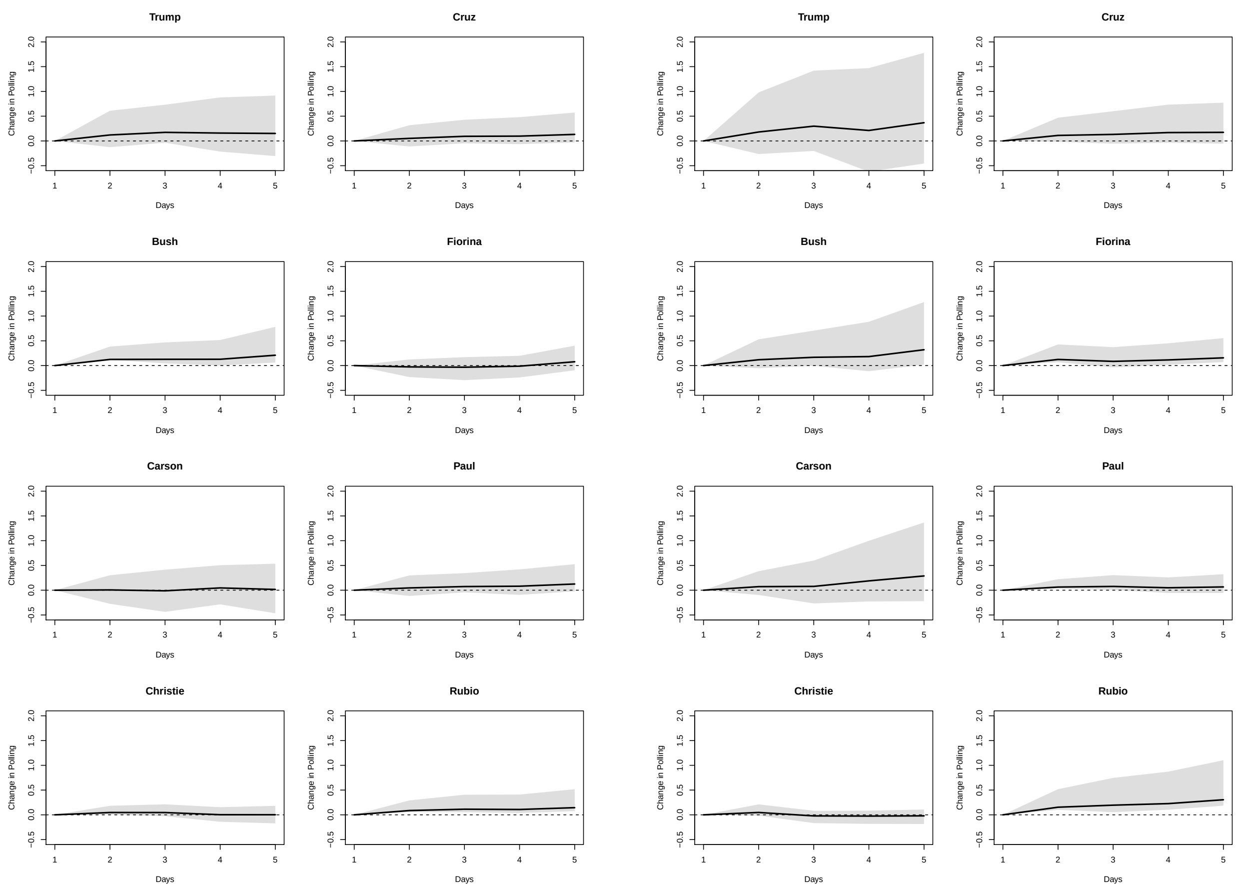

Figure 23: Comparing the IRFs of a shock to Interest on Media presented in the main text (left side) to IRFs with an alternative identification specification.

Figure 24: Interest → Support Robustness

Original | Robust to Polling Aggregation

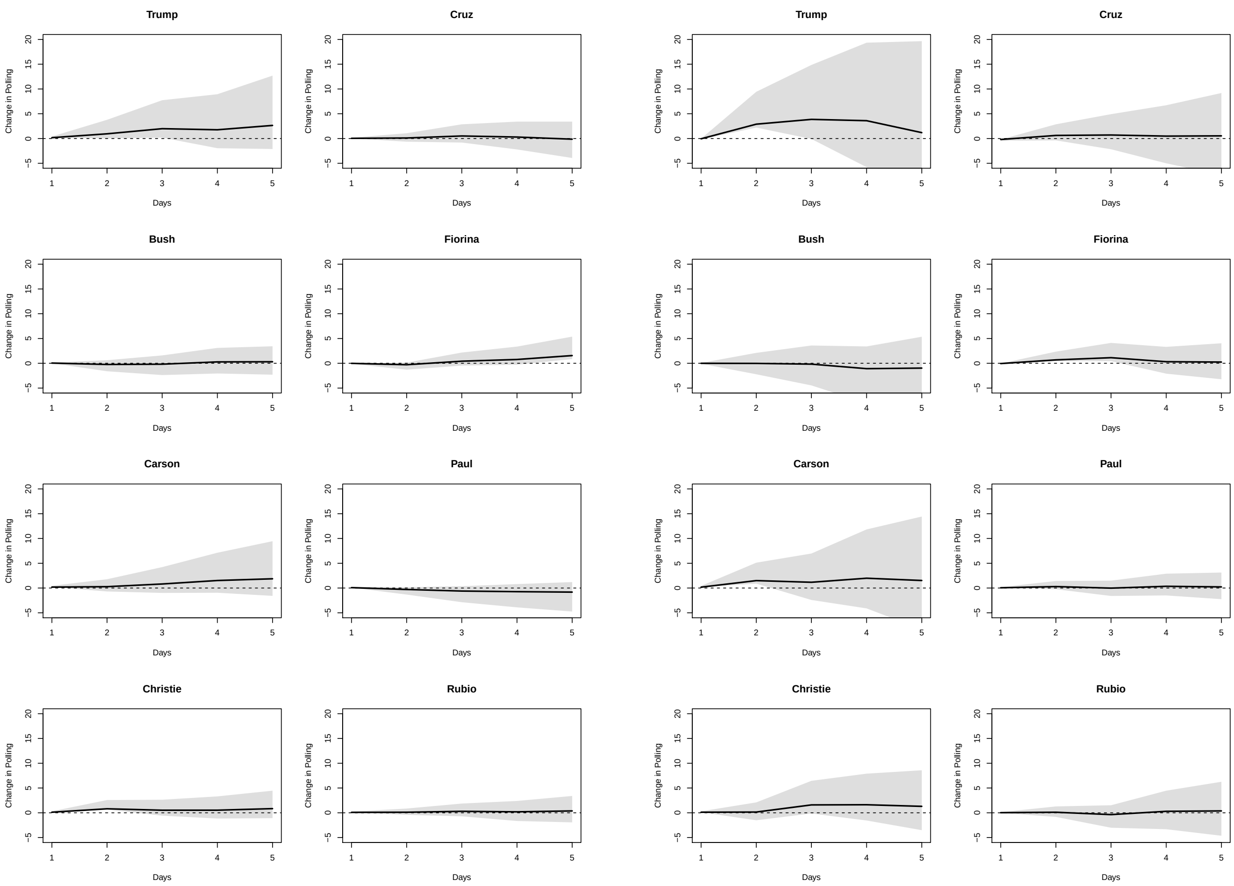

Figure 24: Comparing the IRFs of a shock to Interest on Support presented in the main text (left side) to IRFs with an alternative identification specification.

References

Ashley, Richard A. and Randal J. Verbrugge. 2009. "To difference or not to difference: a Monte Carlo investigation of inference in vector autoregression models." *International Journal of Data Analysis Techniques and Stategies* 1(3):242–274.

Box-Steffensmeier, Janet M., John R. Freeman, Matthew P. Hitt and Jon C. W. Pevehouse. 2014. *Time Series Analysis for the Social Science*. Cambridge University Press.

Brandt, Patrick and John T. Williams. 2007. *Multiple Time Series Models*. SAGE.

Brandt, Patrick T. and John R. Freeman. 2006. "Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting, and Policy Analysis." *Political Analysis* 14(1):1–36.

Brandt, Patrick T. and John R. Freeman. 2009. "Modeling Macro-Political Dynamics." *Political Analysis* 17(2):113–142.

Brandt, Patrick T, Michael Colaresi and John R Freeman. 2008. "The dynamics of reciprocity, accountability, and credibility." *Journal of Conflict Resolution* 52(3):343–374.

Brandt, Patrick and W. Ryan Davis. 2015. "Markov-Switching, Bayesian, Vector Autoregression Models.". Version: 0.9-2.
URL: *https://cran.r-project.org/web/packages/MSBVAR/index.html*

Geweke, John. 1992. Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. In *Bayesian Statistics*, ed. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith. Vol. 4 Oxford pp. 169–194.

Grant, Taylor and Mathew J. Lebo. 2016. "Error Correction Methods with Political Time Series." *Political Analysis* 24(1):3–30.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Wiley Series in Probability and Statistics.

Keele, Luke, Suzanna Linn and Clayton McLaughlin Webb. 2016*a*. "Concluding Comments." *Political Analysis* 24(1):83–86.

Keele, Luke, Suzanna Linn and Clayton McLaughlin Webb. 2016*b*. "Treating Time with All Due Seriousness." *Political Analysis* 24(1):31–41.

Ljung, Greta M and George EP Box. 1978. "On a measure of lack of fit in time series models." *Biometrika* 65(2):297–303.

Mills, Terence C. 1998. "Recent Developments in Modeling Nonstationary Vector Autoregressions." *Journal of Economic Surveys* 12(3):279–312.

Phillips, Peter C. 1995. "Fully Modified Least Squares and Vector Autoregression." *Econometrica* 63.

Phillips, Peter C. 1998. "Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VAR's." *Journal of Econometrics* 83(1-2):21–56.

Sims, Christopher A. and Tao Zha. 1998. "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review* 39(4):949–968.

Toda, Hro Y. and Taku Yamamoto. 1995. "Statistical Inference in Vector Autoregressions with Possible Integrated Processes." *Journal of Econometrics* 66(1-2):225=250.

Waggoner, Daniel F. and Tao Zha. 2003. "Likelihood Preserving Normalization in Multiple Equation Models." *Journal of Econometrics* 114:329–347.

Williams, John T. 1993. "What goes around comes around: Unit root tests and cointegration." *Political Analysis* 4:208–244.