Making Sense of Isolationism:

Foreign Policy Mood as a Multilevel Phenomenon

Supplementary Appendix (Updated: December 14, 2012)

Joshua D. Kertzer

Contents

This document is an online appendix that contains additional information excluded from the manuscript itself due to space limitations. I begin with a more detailed justification of the use of a hierarchical model and a discussion of the model fit criteria employed in the text. Second, since models predicting foreign policy mood are only useful if the measure of foreign policy mood displays convergent validity – that is, it predicts what we would expect it to predict – I present a series of analyses showing the substantive impact of the foreign policy mood measure on a number of specific policy attitudes tapping into both militant and cooperative forms of internationalism. Third, I discuss multiple imputation in slightly greater detail than in the main manuscript, before elaborating further on the "reality constraint" that weighs the impact of subjective economic assessments on foreign policy mood, presenting additional data showing how the variance of subjective economic assessments shrinks during periods of economic turmoil. Fourth, I present further discussion of the surprisingly large impact of personal efficacy, before discussing the algorithm used to estimate the joint hypothesis tests presented in the main document. Finally, I present a number of supplementary analyses using rally events as an alternative measure of the international security environment, and conclude with with a number of alternative model specifications.

1 Adjudicating model fit

In a traditional regression context, likelihood ratio tests could be used to formally test the relative goodness of fit of various model specifications, comparing both the extent to which a hierarchical model offers a higher goodness of fit than a non-hierarchical model, as well as adjudicating relative model fits amongst different subtypes of multilevel models. Since the multilevel logistic models presented here are estimated using a penalized quasi-likelihood function, however, formal likelihood ratio tests are unavailable. Thus, model fits are judged less formally using three types of criteria: an in-

tercept plot used to demonstrate temporal heterogeneity (and hence, the need for a multilevel model), and comparisons of ρ parameters and confidence intervals around the intercept, used to compare the relative performance of different model specifications. Each are discussed in turn.

1.1 Justifying a hierarchical model

First, to determine whether a multilevel model is needed, I plot the intercept estimates from separate logistic regressions for each of the 14 waves in the study in Figure 1. The plot shows evidence of heterogeneity across time: even when controlling for our 13 individual-level factors of interest, the estimates for each year look quite different, as the difference in the intercept estimates between 1982 and 2002 illustrate. Considerable temporal heterogeneity remains; the question becomes how to best account for it.[1]

As discussed in the text, temporal heterogeneity is typically taken into account by a number of different means. For example, in his analysis of support for American global activism, Fordham (2008) addresses temporal heterogeneity in ANES data pooled from 1956-2000 by using year fixed-effects. Although this approach addresses the statistical problem — without controlling for this heterogeneity, coefficient estimates would be biased — it treats temporal heterogeneity as a nuisance to be controlled for, rather than a feature of our data that we should try to explain; it is for this reason that Beck (2001) notes that fixed effects are not theoretical variables. Rather than arguing that responses in 1998 differ from those in 2002 because the former were conducted in 1998 and the latter in 2002 (the type of argument implicit in a year-fixed effects approach), a multilevel approach lets us not only control for heterogeneity, but model it theoretically, turning to theoretical variables like the state of the domestic economy or international security environment to explain the differences instead.

[1] In a traditional regression context, likelihood ratio tests could be used to formally evaluate whether multilevel modeling is necessary, but because the multilevel logistic models are estimated using a penalized-quasi likelihood, these formal tests are unavailable, and the visual test provided by Figure 1 is used instead.

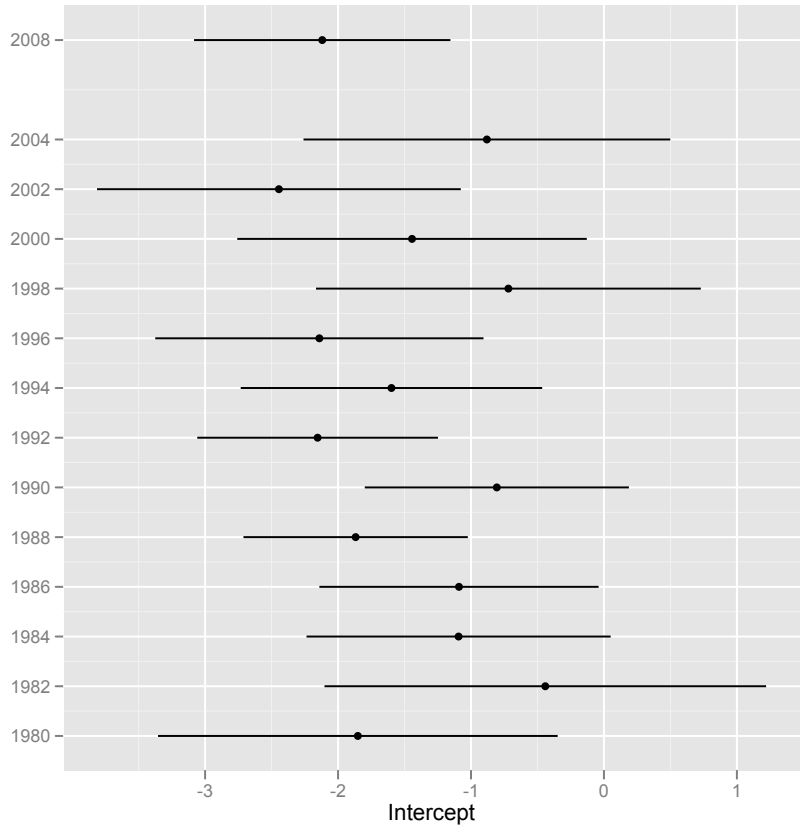


Figure 1: The intercept estimates from separate logistic regressions for each year show signs of heterogeneity, despite the inclusion of the 13 individual-level predictor variables.

Political scientists are relatively familiar with multilevel models — indeed, Kedar and Shively (2005, 2) argue that "all comparative politics is multilevel" — but have tended to exploit their statistical properties only when dealing with actors nested in spatially-differentiated units: students nested in schools (Hill and Rowe, 1996), voters nested in states (Gelman et al., 2007; Keele and Wolak, 2008), survey participants nested in countries (Anderson and Tverdova, 2003; Hutchison and Gibler, 2007), and so on.[2] Ultimately, however, multilevel models are about studying the effects of context (Steenbergen and Jones, 2002; Falleti and Lynch, 2009), and since each survey consists of a different random sample of respondents — rather than panel data in which respondents reappear across each wave of the study — we can treat each biannual NES poll as representing a different level-two unit.

[2] Not all applications of multilevel models by political scientists have been in terms of spatial context, as Bartels's (2009) modeling of justices' choices nested in cases nested in years demonstrates.

As such, we can look at the domestic economic and international security environments in which respondents were situated at the time that they participated in the surveys, and use them to study the impact to which butter and guns, respectively, condition the effect of individual-level determinants of foreign policy views.

1.2 Choosing a functional form: what kind of hierarchical model?

The second place where model fit becomes an issue involves choosing between multiple specifications of hierarchical model. To evaluate the relative goodness of fit of different types of hierarchical models, two types of criteria are used. The first is ρ, the *intraclass correlation*, which indicates how much of the variation in the model is due to differences between groups, rather than differences within them: the better-fitting the model, the lower the ρ parameter should be, as more of the between-group variation is being explained. The second is the confidence interval around the intercept, expressed in predicted probability form: the narrower the confidence interval, the better-fitting the model.

As Steenbergen and Jones (2002) note, multilevel models come in a number of different submodels; although full hierarchical models are employed for most of the analyses in the paper, two additional types of multilevel models are used. The first is a *oneway ANOVA* model (also called a *Random Effects ANOVA*), which contains no predictor variables, and is used simply to partition variance between the model's levels. The level-one and level-two equations can be expressed using equations 1 and 2 respectively, substituting in to produce the multilevel equation 3, in which $i = 1...N$ units are nested into $j = 1...J$ groups, whose intercepts vary:

$$Y_{ij} = \mathrm{B}_{0j} + \epsilon_{ij} \tag{1}$$

$$\mathrm{B}_{0j} = \gamma_{00} + u_{oj} \tag{2}$$

$$Y_{ij} = \gamma_{00} + u_{oj} + \epsilon_{ij} \tag{3}$$

Although a oneway ANOVA is used to establish a baseline ρ and set of confidence intervals about the intercept (the results for which are presented in model 1 of Table 8), its lack of predictor variables makes it less theoretically interesting than a *random intercept* model, which not only allows intercepts between groups to vary, but also includes a set of predictor variables (simplified here as vector Z_j):

$$Y_{ij} = \mathrm{B}_{0j} + \epsilon_{ij} \tag{4}$$

$$\mathrm{B}_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{oj} \tag{5}$$

$$Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + u_{oj} + \epsilon_{ij} \tag{6}$$

A series of random intercept models were estimated as an intermediate step to gauge model fit (see Table 8) but not discussed in the main document. For most of the analyses reported in the text, however, a *full hierarchical model* is used; it's also denoted in the tables as models with a *random coefficient*, in that it not only permits the intercepts to vary across groups (equation 8), but also allows the slopes of at least one of the variables to vary across groups as well (equation 9):

$$Y_{ij} = \mathrm{B}_{0j} + \mathrm{B}_{1j} + \epsilon_{ij} \tag{7}$$

$$\mathrm{B}_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{oj} \tag{8}$$

$$\mathrm{B}_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{ij} \tag{9}$$

$$Y_{ij} = \gamma_{00} + \gamma_{01} Zj + u_{oj} + \gamma_{10} + \gamma_{11} Z_j + u_{ij} + \epsilon_{ij} \tag{10}$$

Table 1: Evaluating improvement in fit across alternative model specifications

Model	ρ	CI on intercept
Baselines		
Oneway ANOVA	0.073	[0.607, 0.823]
Random intercept model	0.058	[0.134, 0.291]
Random coefficient models		
Retro. sociotropic	0.100	[0.113, 0.318]
Pro. sociotropic	0.114	[0.110, 0.334]
State efficacy	0.057	[0.133, 0.285]
Personal efficacy	0.013	[0.170, 0.225]
Party ID	0.048	[0.138, 0.277]
Ideology	0.057	[0.132, 0.284]
Age	0.073	[0.125, 0.298]
Male	0.056	[0.133, 0.283]
Income	0.013	[0.161, 0.230]
Knowledge	0.022	[0.158, 0.250]
Economy salient	0.047	[0.137, 0.274]
Inpartisan	0.035	[0.144, 0.261]]

The gradual improvement in both ρ and the confidence interval about the intercept described in the text is used to settle on a full hierarchical model in which both the intercept and the slope of personal efficacy are allowed to vary across years. Alternative specifications of the model were also estimated in which a random coefficient was placed on each of the other predictor variables; as Table 1 indicates, personal efficacy and income are the two variables for which the inclusion of a random slope most improves model fit. Both models have equally small ρ statistics (0.013), but the former has tighter confidence intervals around the intercept, so the analyses in the main document include a random slope on the former rather than the latter.[3] Importantly, for most of the variables, including a random coefficient either offered no significant improvement in fit (in which ρ barely decreased compared to the baseline), or actually produced a reduced fit. These ρ statistics thus are also substantively interesting, since they imply that the effects of partisanship and political ideology on foreign policy mood can essentially be treated as being constant across the 1980-2008 period, rather

[3] A random coefficient model was also estimated with a random slope on education, but because the model did not converge for 7 of the 10 imputed datasets, its statistics are not reported in Table 1.

than fluctuating across different Presidential administrations or periods of US foreign policy.

2 Convergent validity: what does foreign policy mood predict?

Many scholars of public opinion about foreign policy have rejected a simple unidimensional internationalism-isolationism measure, noting that we can distinguish militant forms of internationalism — which concern attitudes towards the use of force — from cooperative forms of internationalism, which emphasize institutions and international cooperation (Holsti, 1979; Wittkopf, 1986, 1990). Hierarchical models of constraint in foreign policy belief systems (e.g. Hurwitz and Peffley, 1987; Liberman, 2006; Rathbun, 2007) emphasize the role that general attitudes play in constraining more specific ones, such that our generic measure of support for isolationism should also be able to predict support for specific policies that tap into both cooperative and militant internationalism.

Accordingly, a series of OLS and ordered probit models were estimated on a number of the more specific foreign policy questions from the 2008 ANES, examining the extent to which the generic measure of foreign policy mood predicts support for specific policies even when controlling for partisan identification, political ideology, and a series of demographic characteristics. The 2008 wave of the ANES includes a number of questions regarding foreign policy priorities: respondents were presented with a list of foreign policy goals (e.g. fighting terrorism, combatting hunger, and so on) and asked whether each goal was a very important priority, an important priority, or not an important priority. The survey also included questions measuring the extent to which respondents approved or disapproved of the handling of the wars in Iraq and Afghanistan. Table 2 presents the results for those policies associated with cooperative internationalism. Even when controlling for party ID, demographic characteristics, and political ideology, our foreign policy mood variable strongly predicts support for specific policy preferences. Foreign policy extroverts were consistently more likely to de-

scribe the sorts of humanitarian and cooperative policies favored by cooperative internationalists as being very important, and consistently less likely to describe them as being not important: compared to our isolationists, foreign policy extroverts were 9% more likely to endorse strengthening the United Nations as a very important foreign policy goal, 12.0% more likely to say the same about championing hunger, 12.4% more likely to call for promoting human rights, 6.7% more likely to favor stemming nuclear proliferation, 4.4% more likely to advocate for promoting markets, and 14.7% less likely to describe promoting democracy as a not very important priority. Table 3 finds similar results for policy questions tapping into militant internationalism: foreign policy extroverts were 11.5% more likely to advocate fighting terrorism, and although partisan attitudes towards the wars in Iraq and Afghanistan were quite strong by 2008, foreign policy extroverts were still 0.42 and 0.49 points more supportive of the Iraq and Afghan wars than their introverted counterparts. These results suggest that our measure of foreign policy mood displays convergent validity, in that it predicts what we expect it to.

Table 2: Foreign policy mood predicts support for indicators of cooperative internationalism

	Promoting Democracy		Combatting Hunger		Promoting markets	
FP Extroversion	0.39***	(0.065)	0.304***	(0.071)	0.165**	(0.065)
Age	0	(0.002)	0.002	(0.002)	0.004*	(0.002)
Party ID	-0.062	(0.017)	-0.111	(0.018)	-0.023	(0.016)
Male	-0.041	(0.059)	-0.269	(0.064)	0.265***	(0.059)
Education	-0.119	(0.019)	-0.105	(0.021)	-0.027	(0.019)
Ideology	0.105***	(0.023)	-0.042	(0.026)	0.021	(0.023)
τ_1	-0.384	(0.172)	-2.274	(0.194)	-0.532	(0.173)
τ_2	1.176***	(0.174)	-0.841	(0.186)	1.061***	(0.174)

	Human Rights		Strengthening the UN		Nuclear Proliferation	
FP Extroversion	0.466***	(0.066)	0.25***	(0.066)	0.214**	(0.08)
Age	-0.003	(0.002)	-0.001	(0.002)	0.006**	(0.002)
Party ID	-0.061	(0.017)	-0.14	(0.017)	0.029†	(0.021)
Male	0.035	(0.06)	-0.197	(0.06)	-0.007	(0.075)
Education	0.013	(0.02)	-0.103	(0.02)	-0.019	(0.024)
Ideology	-0.042	(0.023)	-0.076	(0.024)	-0.013	(0.029)
τ_1	-0.95	(0.174)	-2.194	(0.181)	-1.190	(0.216)
τ_2	0.614***	(0.174)	-0.779	(0.175)	-0.310	(0.212)

Main entries are ordered probit coefficients. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ †$p < 0.10$

Table 3: Foreign policy mood predicts support for indicators of militant internationalism

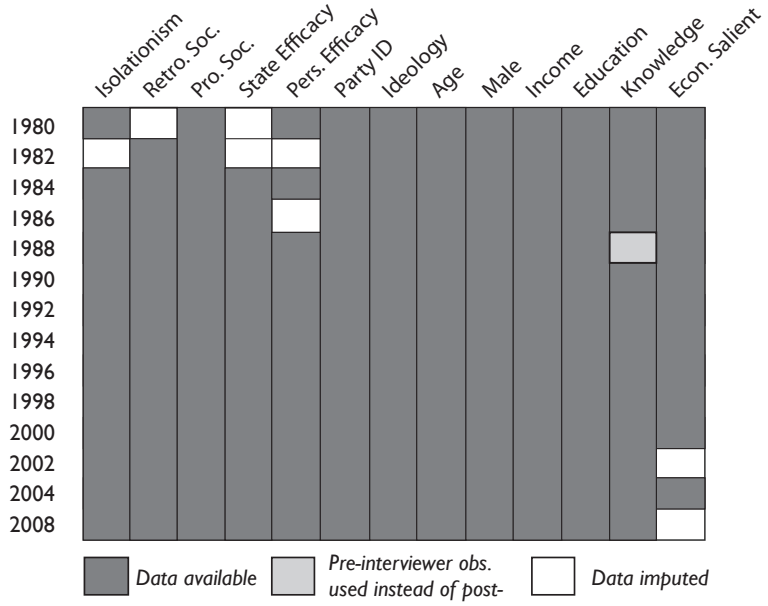
	Fighting Terrorism		Support for Iraq War		Support for Afghan War	
FP Extroversion	0.321***	(0.077)	0.420***	(0.097)	0.492***	(0.095)
Age	0.013***	(0.002)	0.005*	(0.003)	0.001	(0.003)
Party ID	0.036*	(0.021)	0.446***	(0.025)	0.369***	(0.024)
Male	0.037	(0.072)	0.188*	(0.088)	0.151*	(0.086)
Education	-0.078	(0.024)	-0.045	(0.029)	-0.019	(0.028)
Ideology	0.094***	(0.028)	0.179***	(0.034)	0.123***	(0.033)
τ_1	-0.682	(0.21)				
τ_2	0.565**	(0.206)				
Intercept			-0.297	(0.257)	0.318	(0.252)

Main entries are ordered probit coefficients in the first model, OLS coefficients otherwise.

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ †$p < 0.10$

3 Variable coverage and multiple imputation

Figure 2: Micro-level variable coverage.



As discussed in the manuscript and visually depicted by Figure 2 above, not all of the micro-level variables are included in each wave of the ANES data. Rather than employ listwise deletion and throw out large amounts of information, multiple imputation was conducted using Amelia II (Honaker, King and Blackwell, 2009), producing $m = 10$ imputed datasets on which analyses were run individually, their results aggregated, as in Rubin (1987), by:

$$\bar{q} = \frac{1}{m}\sum_{j=1}^{m} q_j \tag{11}$$

$$SE(q) = \sqrt{\frac{1}{m}\sum_{j=1}^{m} SE(q_j)^2 + \sum_{j=1}^{m}\frac{(q_j - \bar{q})^2}{m-1}(1+\frac{1}{m})} \tag{12}$$

where $\bar{q}$ is the point estimate of quantity of interest q, and $SE(q)$ its standard error. For simulated confidence intervals for predicted probabilities, as suggested by King et al. (2001), confidence intervals

for predicted probabilities were simulated by drawing 1000 simulations from a multivariate normal on each of the m datasets.[4] A modified version of the same algorithm is used for the joint hypothesis test probing the interactions in Figure 1 of the main document, to determine the probability that the effects of that predictor X differs at each value of moderator Z. See the discussion beginning on p. 16 for more details.

4 The interaction of micro- and macro-level economic factors

The main text includes a series of analyses where micro-level economic assessments are interacted with macro-level economic conditions, showing how the impact of individual-level subjective economic perceptions on foreign policy mood is in fact conditional on the economic realities on the ground: when economic times are good, positive subjective assessments are associated with greater support for an extroverted foreign policy, but when economic times sour, the extroversion gap disappears, and subjective assessments are no longer significant. The mechanism implicated here is that of a "reality constraint" that kicks in when the economy truly sours: subjective economic assessments have a statistically significant impact on foreign policy mood even when controlling for objective reality, but once objective economic conditions worsen sufficiently, this effect is no longer significant. If this were the case, we would expect subjective economic assessments to be more likely to converge in bad economic times than in good ones.

I provide additional evidence in favor of this relationship using a simple visual test. The Michigan Surveys of Consumers — who produce the index of consumer sentiment (ICS) used as a macro-level economic indicator in the analyses presented in the main text — include a question measuring respondents' retrospective sociotropic assessments — that is, evaluating where current national business

[4] For more details on the specific EM algorithm used, see Honaker, King and Blackwell (2009).

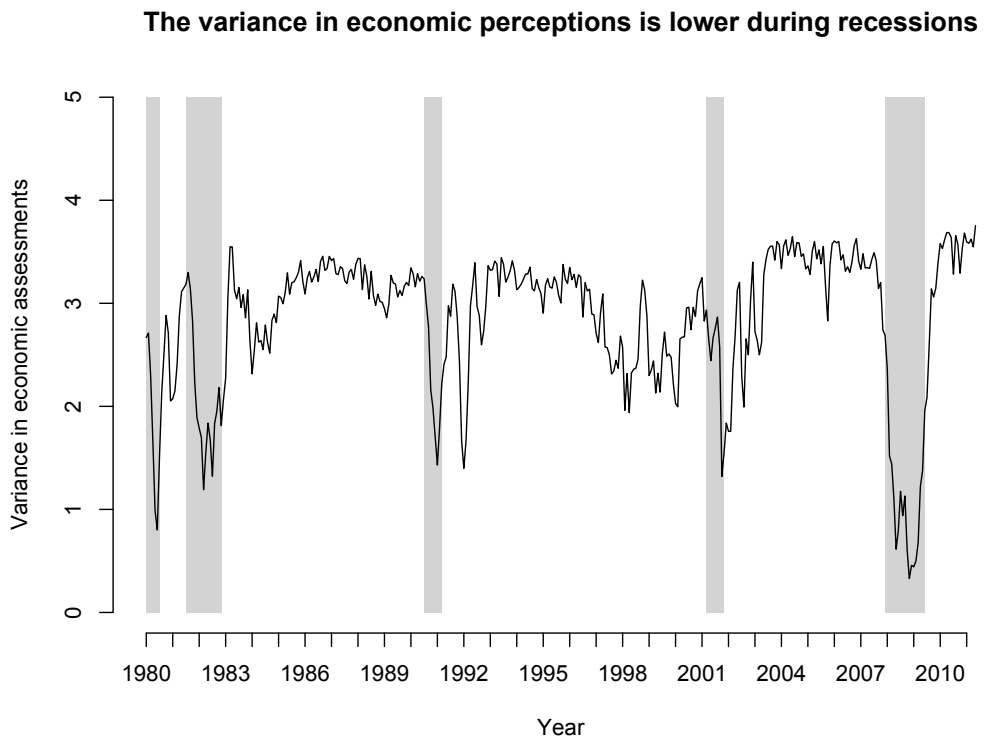


Figure 3: The variance in subjective economic assessments at each time period is measured on the y-axis; the more respondents' assessments converge, the closer the variance is to 0. The grey shaded areas depict periods of recession as determined by the National Bureau on Economic Research (NBER); the plot clearly shows how subjective economic assessments converge during recessions — signs of a reality constraint — but diverge in better economic times, thereby explaining why individual-level economic assessments' impact on foreign policy mood is conditional on macro-level economic conditions.

conditions compared to a year ago are "better", "the same", or "worse". This retrospective sociotropic measure is similar to the one fielded by the ANES that I use as my central measure of subjective economic assessments in the main text, but has the advantage of being fielded monthly rather than biannually, thereby allowing for a more fine-grained time series of subjective economic assessments. Since the degree of convergence in subjective assessments is a question about the *variance* of the distribution rather than of the mean (Braumoeller, 2006), I calculate the variance of subjective economic assessments in each monthly poll from January 1980 to May 2011, plotting them in Figure 3.[5] If every respondent in a poll provides the same assessment about the state of the economy, the

[5] Note that the variance being calculated here is the variance *at each* time period rather than the variance *across time*

variance will be 0; the more subjective assessments diverge, the higher the variance in responses. The shaded grey periods correspond to times when the economy is objectively bad – the periods of recession as determined by the National Bureau of Economic Research (NBER).[6] The plot clearly displays that the variance in individual-level economic assessments is lower during recessions than in periods of economic expansion, suggesting the presence of a powerful reality constraint that kicks in when economic times are most dour, and individual-level perceptions of the state of the economy converge.

periods or volatility typically calculated by time series models that study variance, like Autoregressive Conditional Heteroskedasticity (ARCH) models (Gronke and Brehm, 2002) or Dynamic Conditional Correlation (DCC) models (Lebo and Box-Steffensmeier, 2008).

[6]The NBER's Business Cycle Dating Committee determines the peaks and troughs of the business cycle by looking at a variety of economic indicators, from real GDP and real GDI to measures of total employment, as well as trends in particular economic sectors. In this sense, the NBER's periods of recession are not entirely objective measures — there's no fixed weighting of each indicator, for example — but given the prominent role the NBER plays in adjudicating periods of recession, it can be thought of as a suitable macro-level economic indicator of the state of the US economy. For more information on how the NBER dates business cycles, see http://www.nber.org/cycles/recessions_faq.html.

5 Personal efficacy

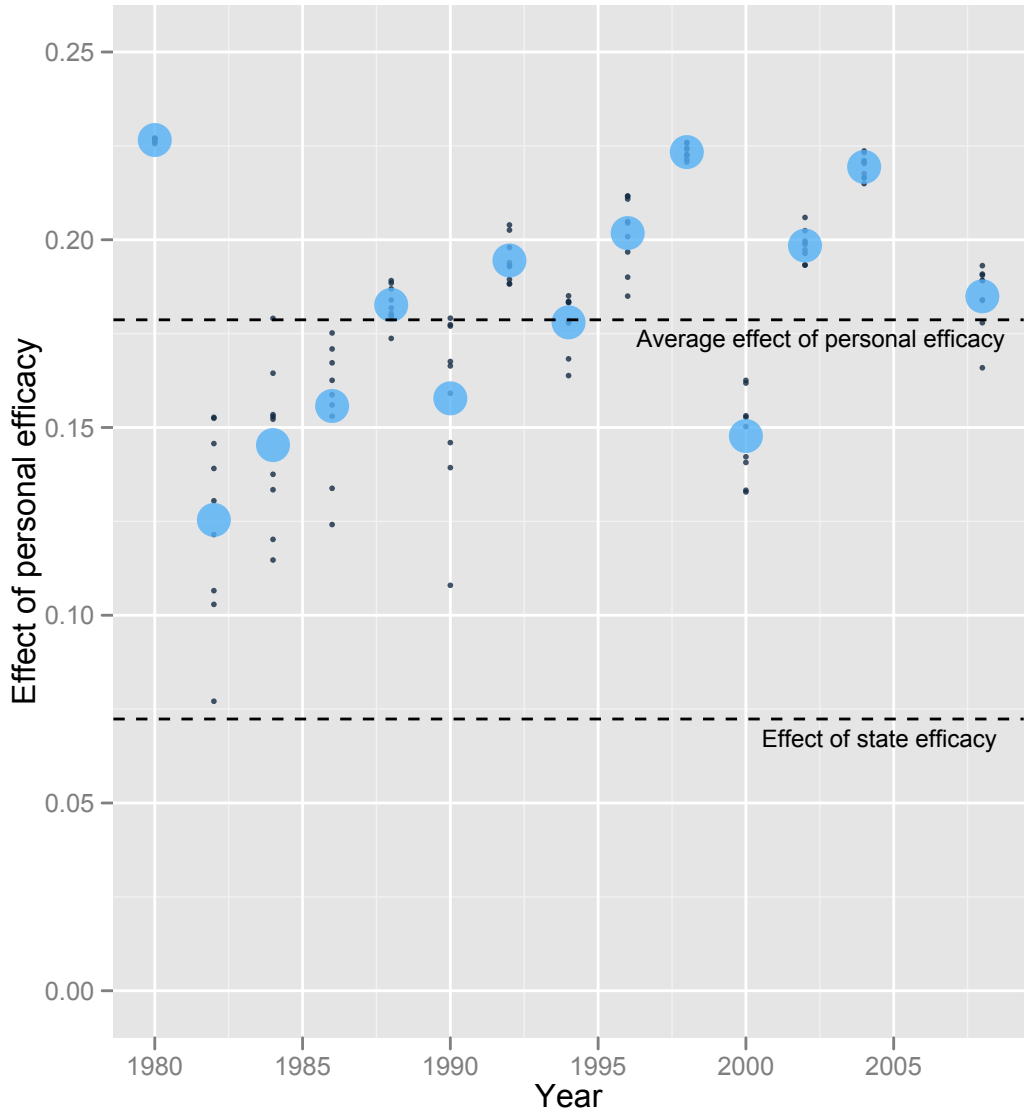


Figure 4: Comparing the effect of personal efficacy on the predicted probability of foreign policy extroversion across years. The large circles represent the change in the predicted probability of foreign policy extroversion produced by moving from the lowest to highest levels of personal efficacy in each year, while the small dots capture the variation in personal efficacy's effect size across imputations, as calculated by extracting the random effects from each of the m parallel models. For example, in 2000, the large circle shows that personal efficacy boosted the predicted probability of foreign policy extroversion by 14.6%, while the small dots show that across each of the m imputed datasets, personal efficacy's effect ranged from a low of 13.3% to a high of 16.3%, thereby showing that the analyses are not particularly sensitive in that year to the imputation process – unlike in 1990, say, where the small dots are more dispersed. Despite the obvious heterogeneity across time (personal efficacy's effect is 53.7% greater in 1998 than in 1984, for example), its effect is always greater than that of state efficacy.

One of the striking results of the individual-level analyses is the strong performance of personal efficacy; individuals high in personal efficacy were almost 18% more likely to advocate an extroverted foreign policy than their low-efficacy counterparts. Notably, as Figure 4 illustrates, although personal efficacy displays heterogeneous effects on foreign policy mood across time, in no year does it have a smaller impact than state efficacy: people who would rather stay home than play a role in their country's political life would also rather their country stayed home than participate in the international political realm, irrespective of whether they thought the country was actually doing quite well at it. Furthermore, although the inclusion of personal efficacy was originally justified in the model out of the possibility that it correlated with state efficacy, against expectations, the two efficacy variables were not correlated with one another ($r = 0.08$). Although beyond the scope of this paper, it is thus worth investigating in future work why personal efficacy has the substantively significant effect that it does on foreign policy mood.

6 Joint hypothesis tests

In Figure 1 of the main text, I employ joint hypothesis tests rather than more conventional confidence intervals, to formally test the hypothesis that the effects of subjective economic assessments on foreign policy mood differ at each value of objective economic conditions. Figure 5, below, replicates the figure from the main text while overlaying 95% confidence intervals. Since the distinction between the confidence intervals and the joint hypothesis test is a subtle one, I elaborate on it further here, illustrating with `R` code and sample output.

Even though both rely on $\alpha = 0.05$, the confidence intervals and joint hypothesis test are reporting fundamentally different things: the CIs depict our uncertainty around each set of point estimates, whereas the joint hypothesis test focuses on the "gap" between the two sets of point estimates. Al-

though this distinction is often overlooked by political scientists, the fact that two pairs of confidence intervals overlap does not necessarily imply that the effects are statistically insignificant from one another. Recall that the probability mass of a pair of confidence intervals is centered around the mean rather than the tails; the area subsumed by the intersection of the tails of two overlapping sets of CIs — which is our quantity of interest here — may thus be quite small. Rather than take the product of the integrals of each of the density distributions — which would assume independence, which is unlikely in this case — I use simulation-based methods. As shown in the R code below (which is a simplified version of the algorithm used to conduct the tests; the full algorithm is included in the replication materials), I simulate 1000 draws from a multivariate normal, given the covariate profile and variance-covariance matrix from each of the $m = 10$ imputed datasets, producing 10,000 simulations in total, as one would normally do when calculating predicted probabilities on data generated by multiple imputation.[7] I then use these 10,000 simulations to calculate predicted probabilities across each combination of the predictor and moderator, and note the transition point on the x-axis beyond which less than 95% of the predictions when economic assessments are positive exceed the predictions when economic assessments are negative. That is to say, for each of the panels of Figure 5(a), I calculate the probability that the thin line is above the thick one, and draw the dashed vertical line to denote the point at which this probability drops below 95%. Note that the joint hypothesis test does not address whether the two sets of effects have the same slope, which is tested instead by the traditional test of an interaction effect: whether the coefficient on the interaction term is statistically significant.

```
##### Simplified joint hypothesis test algorithm (for full code, see replication material) #####
#m: the number of imputed datasets on which the model is estimated
#k: the number of Beta coefficients in the model
#Betas: a k x m matrix of k Beta coefficients from each of the m imputations
```

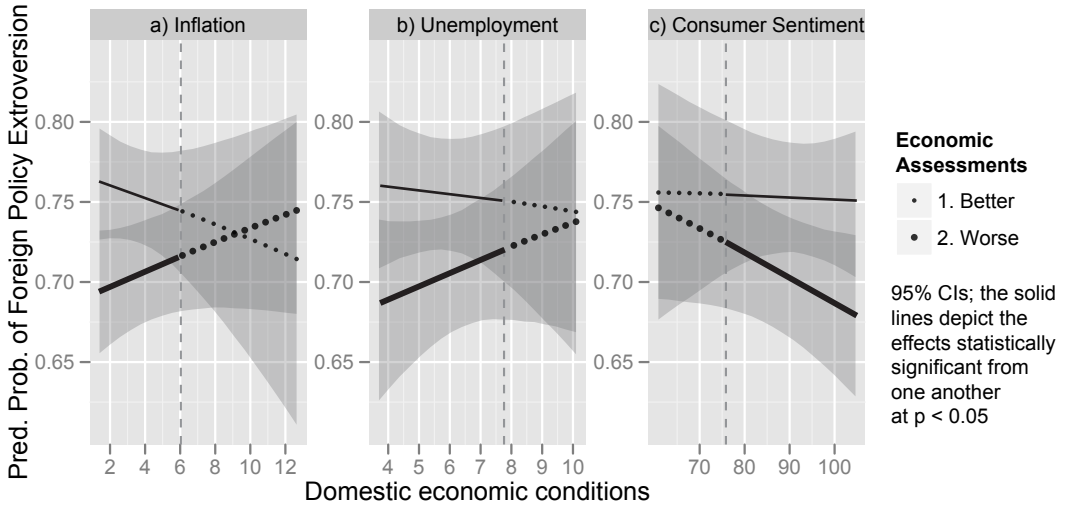
[7] On this point, see section 3 of this supplementary appendix.

```
#Vcovs: a k x k x m array of variance-covariance matrices from each of the m imputations

#Simulate 1000 draws from a multivariate normal given the Betas and variance-covariance
  matrix from each of m imputed datasets, producing 1000*m simulations in total
sim.betas <- matrix(NA, nrow=1000*m, ncol=k)
for (i in 1:m){
  sim.betas[(1000*(i-1) + (1:1000)),] <- mvrnorm(n=1000, mu=Betas[,i], Sigma=Vcovs[,,i])
}

#X: a sequence of values for the focal predictor on the X-axis, ranging from its min to its max
#Z: a sequence of values for our moderator (either low, or high)

#output: an array of size (the length of X)x2, where each row contains the p-value from
   the hypothesis test for that particular value of X, and the columns depict the direction
   of the inequality
output <- array(dim=c(length(X), 2), dimnames=list(round(X,digits=2),
  c("Pr(Zmax > Zmin)", "Pr(Zmin > Zmax)")))

#sim: a matrix of size 1000*m x 2, one row for the results from each simulation, and one
  column for each value of the moderator Z
sim <- matrix(data=NA, nrow=nrow(sim.betas), ncol=2)

#x_val: a vector of length k, representing the covariate profile (in which all continuous
  variables have been set to their means, ordinal variables to their medians, and dummy
  variables to their modes)
#loc.i: a vector of length 3 that identifies the location of the focal predictor X, moderator Z,
   and interaction XZ in the list of k coefficients

#Loops through each combination of X and Z, and saves the proportion of the predictions
  when Z is high that are greater than when Z is low, and vice versa
for (i in 1:length(X)){
  x_val[loc.i[1]] <- X[i]
    for (j in 1:length(Z)){
      x_val[loc.i[2]] <- Z[j]
      x_val[loc.i[3]] <- X[i]*Z[j]
      sim[,j] <- plogis(sim.betas %*% x_val)
    }

  output[i,] <- c(sum(sim[,2] > sim[,1], na.rm=TRUE)/nrow(sim),
    sum(sim[,1] > sim[,2], na.rm=TRUE)/nrow(sim))
}
output
```

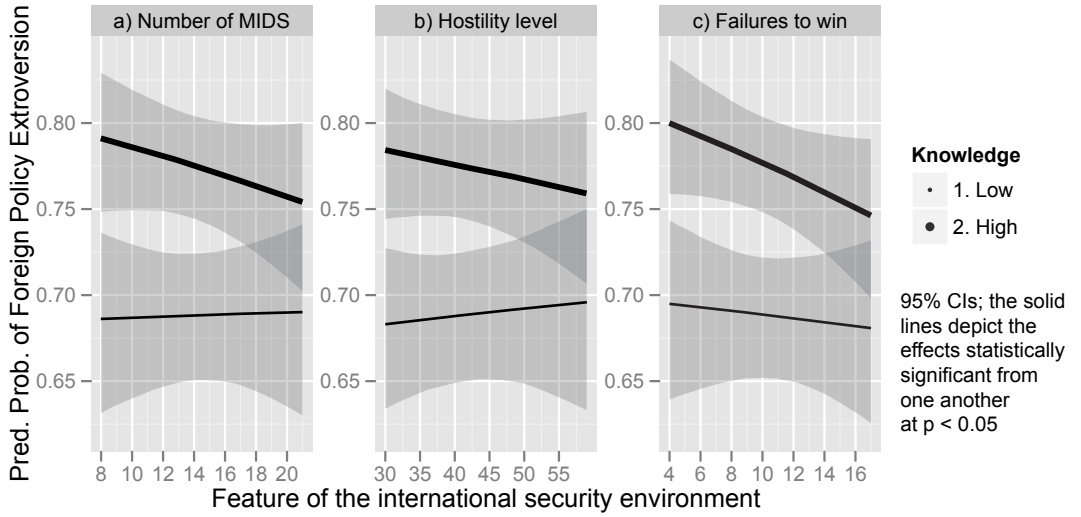
Thus, the output table below shows that for model 2 in Table 1 of the main document (plotted in Figure 5(a)), we can reject the null hypothesis that economic assessments have no effect on foreign policy mood with $p < 0.05$ up until inflation sits at slightly over 6%. At this point, the "reality constraint" kicks in, and subjective assessments appear to no longer have a significant effect.

```
> output
       Pr(Zmax > Zmin)   Pr(Zmin > Zmax)
1.4             0.0000           1.0000
1.87            0.0000           1.0000
2.34            0.0000           1.0000
2.81            0.0000           1.0000
3.27            0.0000           1.0000
3.74            0.0000           1.0000
4.21            0.0001           0.9999
4.68            0.0008           0.9992
5.15            0.0028           0.9972
5.62            0.0125           0.9875
6.08            0.0424           0.9576
6.55            0.0963           0.9037
7.02            0.1671           0.8329
7.49            0.2462           0.7538
7.96            0.3298           0.6702
8.43            0.4035           0.5965
8.89            0.4627           0.5373
9.36            0.5171           0.4829
9.83            0.5624           0.4376
10.3            0.6027           0.3973
10.77           0.6354           0.3646
11.24           0.6635           0.3365
11.7            0.6865           0.3135
12.17           0.7089           0.2911
12.64           0.7306           0.2694
```

Figure 5: Macro-level determinants of foreign policy mood, with 95% CIs



(a) A replication of Figure 1(a) from the main document, but with 95% confidence intervals displayed in addition to the joint hypothesis test highlighted in the original version of the figure. The dashed vertical line illustrates the points along the x-axes for which we can no longer be certain with 95% probability that foreign policy mood differs across levels of economic assessments. Thus, when (i) inflation and (ii) unemployment rates are low or (iii) consumer confidence is high, subjective readings about the economy have a statistically significant effect on foreign policy mood. Once these objective economic indicators pass a certain threshold, though, a "reality constraint" kicks in, and subjective economic evaluations lose their impact



(b) A replication of Figure 1(b) from the main document, but with 95% confidence intervals displayed in addition to the joint hypothesis test highlighted in the original version of the figure. The absence of the dashed vertical line shows that we can be certain with over 95% probability that the effects of knowledge differ across the full range of values on the x-axis. The results show that low-knowledge individuals are generally unaffected by the international security environment, while higher-knowledge counterparts are less likely to advocate an extroverted foreign policy as the environment becomes more conflictual, as measured both by (i) the number of militarized disputes the US is involved in during each period, and (ii) the level of escalation of the disputes. Information about (iii) failures to win, however, trickles down even to those low in political knowledge, although they react less sharply to these defeats and stalemates than high-knowledge individuals.

7 Rally events

One of the central findings of the macro-level analyses was that the public's foreign policy mood becomes more introverted when the US is engaged in more militarized international disputes: rather than responding to a belligerent international environment by standing firm, the public — particularly the segment high in political knowledge — is more likely to turn inward. This image of the public is consistent with the literature emphasizing the extent to which democratic publics are sensitive to the costs of war, used to explain why the advantages of democracy in battle decline over time (Bennett and Stam, 1998; Reiter and Stam, 2002). However, research on the rally around the flag effect paints the opposite image, in which the public responds to military crises by becoming *more* supportive of the President's foreign policy agenda (Mueller, 1973; Parker, 1995; Hetherington and Nelson, 2003; Kam and Ramos, 2008). The results in the main analyses thus seem to support the cost sensitivity account over the rally model, but there are two reasons why the analyses might be stacked in the former's favor. First, the two theories have different time scales — rallies occur immediately after a crisis and dissipate soon thereafter, whereas cost sensitivity manifests itself in the long-run — and the time lag in the MID analyses might therefore make rallying harder to observe. Second, rally events are, by definition, highly salient episodes that receive considerable media coverage, whereas many of the MIDS in the data are less prominent, hence why their effect on foreign policy mood appears largely limited to the knowledgeable.

As an alternative test, I therefore turn to Newman and Forcehimes's (2010) list of rally events; the advantage of this particular list of rally events is that the authors ensure comprehensiveness by compiling lists of candidate events from previous rally studies, but also ensure selectivity by requiring that the event have received front-page coverage in the *New York Times* at least 3 times in a month, thereby eliminating the less prominent events. Newman and Forcehimes divide rally events into eight categories: domestic, international, diplomatic, and personal – each of which can be positive or neg-

ative. For the analyses below, however, I focus solely on positive and negative international events, for two reasons. First, my interest is in using rally events as an alternative measure of the international security environment; personal events involving the President (such as President Reagan's colon surgery in July 1985, or the breaking of the Monica Lewinsky scandal in January 1998) or events in domestic politics (e.g. the New Hampshire primary in 1992, the budget deadlock in 1995) thus are less relevant here, even though they may affect Presidential approval ratings. I exclude diplomatic events because on the whole, they are less likely to be salient than the type of major military developments that typically constitute international rally events. Second, I include both positive and negative international events because the meaning of "positive" and "negative" in a rally context is somewhat counterintuitive, since events like bombings of US embassies in Kenya and Tanzania are coded as positive, because rally theories suggest that crises and military conflicts should bolster Presidential popularity.[8]

The other important issue with rally effects is the question of time lags: given that rallies are typically short-lived phenomena, it is unclear that an event in January 1980 should affect the public's foreign policy mood in the Autumn of 1982. Accordingly, in the analyses that follow, I estimate the impact of rally effects in three different ways: first, focusing on the number of rally events that have taken place in the previous year (Table 4), second, on the number of rally events that have taken place in the summer before the ANES wave was fielded (defined as the period between May and the first week of September) (Table 5), and finally, an alternative measure of summer rally events that also includes 9/11, which otherwise would be excluded from the analyses because it took place in the fall of the preceding year (Table 6). Additionally, Table 7 replicates Table 6, but also controlling for the same MID indicators used in the security analyses in the main document. The results in Table 4 show that a one-year time span is too long for rally events to impact foreign policy mood,

[8]Indeed, the only international events coded as negative in Newman and Forcehimes' typology in the period under investigation are the Operation Eagle Claw in May 1980, the Bitburg controversy in 1985, and revelation about the abuses at Abu Ghraib in May 2004.

but Tables 5-6 suggest that rally events significantly affect foreign policy mood when the time span is shortened to the previous summer. More important, though, are two findings. First, since rally events are by definition highly salient and receive large amounts of media coverage, unlike with the MID data, there are no significant interactions between rally events and knowledge. These results show that the conclusions we draw about the relationship between political knowledge, the security environment, and foreign policy mood thus depends on how restrictive our measures of the security environment are. That said, the results from Table 7 show that re-estimating the security analyses from the main document while also controlling for rally events does not significantly alter the results. Second, rally events do not display a linear relationship with foreign policy mood, but seem to display a highly significant quadratic relationship instead: as Figure 6 shows, a single rally event the preceding summer corresponds to a 1.6% boost in foreign policy extroversion, but the presence of a second event actually lowers foreign policy mood by 1.9%. These results should be considered preliminary, both because of the low number of rally events taking place in each time period, and also because of its somewhat unexpected nonlinear functional form, which thus represents a potentially interesting puzzle for future research. It is nonetheless noteworthy that the above results paint a similar picture of the public as found in the main analyses: the more active the US foreign policy agenda, the more skeptical the public is about it.

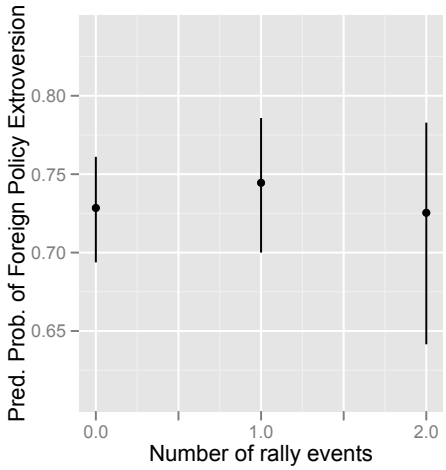


Figure 6: Rally events display a significant quadratic relationship with foreign policy mood, as shown in this plot of the second column of Table 6. A single rally event in the preceding summer (or, in the case of 2002, the 9/11 attacks), corresponds with a 1.6% increase in foreign policy extroversion, but the appearance of a second event is associated with a 1.9% drop. Simulations (see the discussion of the joint hypothesis test beginning on p.16) suggest we can be reasonably confident that foreign policy mood is more extroverted if the public has recently experienced 1 rally event rather than 0 ($p < 0.025$), and more introverted if the public has recently experienced 2 rally events rather than 1 ($p < 0.004$), although it is less clear whether foreign policy mood is in fact more introverted given 2 rally events rather than 0 ($p < 0.096$).

Table 4: Rally events and foreign policy mood (I)

	Rally events in previous year			
	(1)	(2)	(3)	(4)
Micro-level				
(Intercept)	-1.512***	-1.472***	-1.572***	-1.54***
	(0.258)	(0.261)	(0.259)	(0.265)
Retro. Sociotropic	-0.131***	-0.131***	-0.129***	-0.13***
	(0.031)	(0.031)	(0.031)	(0.031)
Pro. Sociotropic	(0.015)	(0.016)	(0.015)	(0.015)
	(0.030)	(0.030)	(0.030)	(0.030)
State Efficacy	0.158***	0.158***	0.158***	0.158***
	(0.025)	(0.025)	(0.025)	(0.025)
Personal Efficacy	0.474***	0.474***	0.473***	0.473***
	(0.060)	(0.059)	(0.060)	(0.059)
Party ID	0.027	0.027	0.027	0.027
	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.010	0.010	0.010	0.010
	(0.017)	(0.017)	(0.017)	(0.017)
Age	(0.011)	(0.011)	(0.011)	(0.011)
	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.072*	0.072*	0.072*	0.072*
	(0.038)	(0.038)	(0.038)	(0.038)
Income	0.105***	0.105***	0.105***	0.105***
	(0.019)	(0.019)	(0.019)	(0.019)
Education	0.166***	0.166***	0.166***	0.166***
	(0.014)	(0.014)	(0.014)	(0.014)
Knowledge	0.224***	0.204***	0.224***	0.206***
	(0.029)	(0.040)	(0.029)	(0.042)
Economy Salient	0.024	0.024	0.025	0.024
	(0.042)	(0.042)	(0.042)	(0.042)
Inpartisan	0.135***	0.136***	0.138***	0.138***
	(0.037)	(0.037)	(0.037)	(0.037)
Macro-level				
Rallies	0.023	(0.017)	0.220	0.199
	(0.061)	(0.077)	(0.189)	(0.267)
Rallies x Knowledge		0.024		0.012
		(0.029)		(0.098)
Rallies^2			(0.103)	(0.113)
			(0.096)	(0.134)
Rallies^2 x Knowledge				0.006
				(0.050)
Unemployment	0.019	0.019	0.023	0.023
	(0.039)	(0.039)	(0.039)	(0.039)
Random effects				
Intercept	(0.075)	(0.083)	(0.081)	(0.080)
95% CI on Intercept	[0.16,0.203]	[0.163,0.212]	[0.15,0.196]	[0.155,0.201]
Personal Efficacy	(0.171)	(0.169)	(0.172)	(0.170)
95% CI on Efficacy	[0.535,0.692]	[0.536,0.691]	[0.534,0.692]	[0.535,0.691]
σ^2	0.994	0.994	0.994	0.994
ρ	0.010	0.011	0.011	0.011
N	23290	23290	23290	23290
Years	13	13	13	13

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

Table 5: Rally events and foreign policy mood (II)

	Rally events in previous summer			
	(5)	(6)	(7)	(8)
Micro-level				
(Intercept)	-1.481***	-1.557***	-1.458***	-1.524***
	(0.274)	(0.241)	(0.277)	(0.243)
Retro. Sociotropic	-0.131***	-0.131***	-0.131***	-0.131***
	(0.031)	(0.031)	(0.031)	(0.031)
Pro. Sociotropic	(0.015)	(0.016)	(0.015)	(0.016)
	(0.030)	(0.030)	(0.030)	(0.030)
State Efficacy	0.158***	0.161***	0.158***	0.161***
	(0.025)	(0.026)	(0.025)	(0.026)
Personal Efficacy	0.475***	0.479***	0.474***	0.478***
	(0.060)	(0.060)	(0.059)	(0.059)
Party ID	0.027	0.027	0.027	0.027
	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.010	0.012	0.010	0.012
	(0.017)	(0.017)	(0.017)	(0.017)
Age	(0.011)	(0.010)	(0.011)	(0.010)
	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.072	0.074	0.072	0.074
	(0.038)	(0.038)	(0.038)	(0.038)
Income	0.105***	0.104***	0.105***	0.104***
	(0.019)	(0.019)	(0.019)	(0.019)
Education	0.166***	0.168***	0.166***	0.168***
	(0.014)	(0.015)	(0.014)	(0.015)
Knowledge	0.224***	0.224***	0.209***	0.205***
	(0.029)	(0.028)	(0.036)	(0.035)
Economy Salient	0.024	0.018	0.024	0.018
	(0.042)	(0.043)	(0.042)	(0.043)
Inpartisan	0.136***	0.136***	0.136***	0.137***
	(0.037)	(0.037)	(0.037)	(0.037)
Macro-level				
Rallies	(0.008)	0.435†	(0.062)	0.326†
	(0.081)	(0.161)	(0.095)	(0.229)
Rallies x Knowledge			0.030	0.068
			(0.035)	(0.096)
Rallies^2		-0.272†		-0.239†
		(0.091)		(0.134)
Rallies^2 x Knowledge				(0.021)
				(0.057)
Unemployment	0.018	0.018	0.019	0.017
	(0.042)	(0.038)	(0.041)	(0.038)
Random effects				
Intercept	(0.074)	(0.074)	(0.060)	(0.071)
95% CI on Intercept	[0.164,0.208]	[0.154,0.196]	[0.171,0.208]	[0.159,0.2]
Personal Efficacy	(0.173)	(0.170)	(0.169)	(0.166)
95% CI on Efficacy	[0.534,0.693]	[0.537,0.692]	[0.536,0.691]	[0.538,0.691]
σ^2	0.994	0.994	0.994	0.994
ρ	0.01	0.007	0.008	0.007
N	23290	23290	23290	23290
Years	13	13	13	13

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

Table 6: Rally events and foreign policy mood (III)

	Rally events in previous summer (and 9/11)			
	(9)	(10)	(11)	(12)
Micro-level				
(Intercept)	-1.49***	-1.577***	-1.451***	-1.537***
	(0.273)	(0.241)	(0.278)	(0.242)
Retro. Sociotropic	-0.131***	-0.133***	-0.131***	-0.134***
	(0.031)	(0.031)	(0.031)	(0.031)
Pro. Sociotropic	(0.015)	(0.016)	(0.015)	(0.016)
	(0.030)	(0.030)	(0.030)	(0.030)
State Efficacy	0.158***	0.161***	0.158***	0.161***
	(0.025)	(0.026)	(0.025)	(0.026)
Personal Efficacy	0.475***	0.478***	0.473***	0.477***
	(0.060)	(0.061)	(0.059)	(0.060)
Party ID	0.027	0.027	0.027	0.027
	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.010	0.012	0.010	0.012
	(0.017)	(0.017)	(0.017)	(0.017)
Age	(0.011)	(0.010)	(0.011)	(0.010)
	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.072	0.074	0.072	0.073
	(0.038)	(0.038)	(0.038)	(0.038)
Income	0.105***	0.104***	0.105***	0.104***
	(0.019)	(0.019)	(0.019)	(0.019)
Education	0.166***	0.168***	0.166***	0.168***
	(0.014)	(0.015)	(0.014)	(0.015)
Knowledge	0.224***	0.223***	0.205***	0.199***
	(0.029)	(0.028)	(0.037)	(0.037)
Economy Salient	0.024	0.018	0.025	0.018
	(0.042)	(0.043)	(0.042)	(0.044)
Inpartisan	0.135***	0.137***	0.136***	0.137***
	(0.037)	(0.037)	(0.037)	(0.037)
Macro-level				
Rallies	(0.005)	0.437†	(0.066)	0.313†
	(0.084)	(0.159)	(0.101)	(0.225)
Rallies x Knowledge			0.036	0.077
			(0.038)	(0.094)
Rallies^2		-0.268†		-0.23†
		(0.088)		(0.131)
Rallies^2 x Knowledge				(0.024)
				(0.055)
Unemployment	0.019	0.020	0.019	0.020
	(0.041)	(0.038)	(0.040)	(0.038)
Random effects				
Intercept	(0.064)	(0.079)	(0.062)	(0.075)
95% CI on Intercept	[0.166,0.204]	[0.15,0.194]	[0.172,0.209]	[0.157,0.199]
Personal Efficacy	(0.172)	(0.170)	(0.167)	(0.166)
95% CI on Efficacy	[0.535,0.692]	[0.536,0.692]	[0.536,0.69]	[0.538,0.69]
σ^2	0.994	0.994	0.994	0.994
ρ	0.009	0.008	0.008	0.008
N	23290	23290	23290	23290
Years	13	13	13	13

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

Table 7: Rally events in previous summer (and 9/11) and MID characteristics (IV)

	(13)	(14)	(15)	(16)	(17)	(18)
Intercept	-1.681***	-1.799***	-1.769***	-1.934***	-1.56***	-1.631***
	(0.412)	(0.369)	(0.512)	(0.450)	(0.361)	(0.308)
Retro. Sociotropic	-0.129***	-0.131***	-0.129***	-0.13***	-0.127***	-0.133***
	(0.032)	(0.032)	(0.033)	(0.032)	(0.032)	(0.032)
Pro. Sociotropic	-0.006	-0.007	-0.006	-0.007	-0.005	-0.007
	(0.032)	(0.031)	(0.032)	(0.031)	(0.031)	(0.031)
State Efficacy	0.157***	0.160***	0.157***	0.161***	0.161***	0.160***
	(0.025)	(0.027)	(0.025)	(0.027)	(0.027)	(0.027)
Personal Efficacy	0.467***	0.471***	0.466***	0.469***	0.472***	0.472***
	(0.063)	(0.064)	(0.063)	(0.064)	(0.063)	(0.063)
Party ID	0.025	0.025	0.025	0.025	0.025	0.025
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.013	0.014	0.013	0.014	0.014	0.014
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Age	-0.012	-0.011	-0.012	-0.011	-0.012	-0.011
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.069	0.071	0.069	0.071	0.071	0.071
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Income	0.101***	0.100***	0.101***	0.100***	0.100***	0.100***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Education	0.168***	0.170***	0.168***	0.170***	0.169***	0.170***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Knowledge	0.346***	0.341***	0.366	0.363	0.319***	0.323***
	(0.104)	(0.105)	(0.127)	(0.126)	(0.080)	(0.080)
Economy Salient	0.022	0.015	0.022	0.016	0.018	0.015
	(0.042)	(0.045)	(0.042)	(0.045)	(0.045)	(0.045)
Inpartisan	0.139***	0.140***	0.139***	0.141***	0.139***	0.141***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Unemployment	0.022	0.023	0.023	0.026	0.015	0.02
	(0.040)	(0.038)	(0.043)	(0.040)	(0.041)	(0.035)
MIDS	0.01	0.013				
	(0.020)	(0.018)				
MIDS x Know.	-0.009	-0.009				
	(0.007)	(0.007)				
Hostility levels			0.005	0.007		
			(0.009)	(0.008)		
Hostility x Know.			-0.003	-0.003		
			(0.003)	(0.003)		
Failures to Win					0.004	0.004
					-0.02	-0.018
Failures x Know					-0.009	-0.01†
					-0.007	-0.007
Rallies	-0.01	0.381†	-0.006	0.401†	-0.02	0.364†
	(0.080)	(0.162)	(0.085)	(0.169)	(0.079)	(0.149)
Rallies^2		-0.236†		-0.244†		-0.232†
		(0.088)		(0.089)		(0.081)
Random effects						
Intercept	(0.093)	(0.089)	(0.089)	(0.083)	(0.094)	(0.105)
Personal Efficacy	(0.176)	(0.174)	(0.175)	(0.173)	(0.168)	(0.170)
σ^2	0.995	0.995	0.995	0.995	0.995	0.995
ρ	0.012	0.009	0.012	0.008	0.013	0.012
N	22078	22078	22078	22078	22078	22078
Years	12	12	12	12	12	12

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

8 Additional model specifications

A number of additional model specifications are presented in Tables 8 - 13 below. Table 8 extends the individual-level analyses from the main document, in several stages. First, to establish a baseline with which to compare subsequent models, it presents the results from the one-way Analysis of Variance (ANOVA) with year random-effects discussed in the text, to analyze how much variation in foreign policy mood is due to differences over time rather than across individuals. The second column in Table 8 presents a random intercept model (RIM), in which all of the individual-level predictors from the main analyses are included in the model, but without a random slope on personal efficacy; the narrower confidence intervals around the intercept and lower ρ statistic show that the RIM fits the data better than the ANOVA, although still not as well as the random coefficient model used in the main text. The third column replicates the second column, but using personal economic assessments rather than sociotropic ones, while the fourth column combines both sociotropic and personal economic assessments into a single model. The personal economic assessments are not statistically significant at the $p < 0.001$ level, and their substantive effects are limited compared to retrospective sociotropic evaluations. In the combined RIM, a change in retrospective sociotropic evaluations from most to least favorable is associated with a 4.2% decrease in the probability of support for an outward-looking foreign policy, while the same change in retrospective and prospective personal evaluations produces a 1.6% and 1.7% decline, respectively. Thus, sociotropic economic evaluations seem to better predict foreign policy mood than personal ones. Finally, the fifth column in Table 8 adds a random slope to personal efficacy, with the low ρ statistic showing that even when controlling for personal economic assessments, the inclusion of a random slope on personal efficacy greatly improves model fit.

Table 9 depicts a series of models in which inflation, unemployment, and consumer sentiment are interacted with political knowledge; none of the main effects are significant apart from consumer sentiment (although the sign appears to be in the wrong direction), and unemployment does have a

significant interaction effect with knowledge, albeit less significant than the interaction effects between these objective economic indicators and retrospective sociotropic assessments discussed in the manuscript.[9] Tables 10 and 11 replicate the contextual security analyses, but testing for significant main effects of each of the three main security variables, as well as for quadratic effects. In contrast to the significant interaction effects between aspects of the security environment and political knowledge reported in the manuscript, we find that the security environment exerts neither significant main effects, nor quadratic effects. Finally, Tables 12-13 replicate the security analyses from the main text, but controlling for inflation and consumer sentiment, respectively, rather than unemployment as in the main set of models.[10] As the results show, regardless of which economic indicator is being controlled for, the results from the main analyses hold.

[9] Additional analyses show the significance of consumer sentiment as a main effect seems to be model-dependent, based on whether the composite security indicator is also included in the model, which therefore suggests the significant main effect merits little confidence.

[10] Unlike in Table 1 in the main document, in which a composite measure was used to control for security conditions, it is inappropriate to include a composite indicator of the domestic economy, since inflation and unemployment are not intercorrelated, so the reliability of the economic scale is low ($\alpha = 0.300$).

Table 8: Individual-level determinants of foreign policy mood: personal economic assessments

	Oneway ANOVA (1)	RIM Sociotropic (2)	RIM Personal (3)	RIM Combined (4)	RCM Combined (5)
Micro-level					
Intercept	0.985***	-1.402***	-1.586***	-1.287***	-1.296***
	(0.077)	(0.174)	(0.156)	(0.180)	(0.142)
Retro. Sociotropic		-0.12***		-0.108***	-0.109***
		(0.030)		(0.030)	(0.030)
Pro. Sociotropic		-0.024		-0.016	-0.015
		(0.030)		(0.031)	(0.032)
Retro. Personal			-0.056	-0.040	-0.040
			(0.021)	(0.021)	(0.021)
Pro. Personal			-0.056	-0.044	-0.048
			(0.032)	(0.033)	(0.032)
State Efficacy		0.165***	0.178***	0.161***	0.16***
		(0.026)	(0.025)	(0.026)	(0.024)
Personal Efficacy		0.464***	0.464***	0.459***	0.475***
		(0.036)	(0.036)	(0.037)	(0.055)
Party ID		0.028***	0.028***	0.028***	0.028***
		(0.009)	(0.009)	(0.009)	(0.009)
Ideology		0.012	0.012	0.012	0.011
		(0.016)	(0.016)	(0.016)	(0.017)
Age		-0.005	0.004	0.002	0.001
		(0.011)	(0.012)	(0.012)	(0.012)
Male		0.037	0.044	0.035	0.034
		(0.036)	(0.036)	(0.035)	(0.036)
Income		0.100***	0.097***	0.098***	0.099***
		(0.018)	(0.018)	(0.018)	(0.019)
Education		0.172***	0.173***	0.172***	0.171***
		(0.014)	(0.014)	(0.014)	(0.014)
Knowledge		0.21***	0.211***	0.21***	0.211***
		(0.026)	(0.026)	(0.026)	(0.026)
Economy Salient		0.020	0.016	0.021	0.027
		(0.043)	(0.043)	(0.043)	(0.042)
Inpartisan		0.136***	0.153***	0.132***	0.133***
		(0.036)	(0.036)	(0.036)	(0.036)
Random effects					
(Intercept)	(0.281)	(0.241)	(0.235)	(0.241)	(0.089)
95% CI on Intercept	[0.607,0.823]	[0.133,0.283]	[0.114,0.245]	[0.147,0.307]	[0.187,0.246]
Personal Efficacy					(0.161)
95% CI on Efficacy					[0.54,0.688]
σ^2	0.999	0.993	0.994	0.993	0.994
ρ	0.073	0.055	0.053	0.055	0.013
N	25612	25612	25612	25612	25612
Years	14	14	14	14	14

Micro-level: *** $p < 0.001$

Table 9: Contextual economic determinants: main effects and knowledge

	Inflation (1)	Inflation (2)	Unem. (3)	Unem. (4)	ICS (5)	ICS (6)
Micro-level						
Intercept	-1.416***	-1.408***	-1.664***	-1.93***	-1.105**	-1.016*
	(0.336)	(0.340)	(0.381)	(0.372)	(0.470)	(0.593)
Retro. Sociotropic	-0.126***	-0.126***	-0.125***	-0.124***	-0.128***	-0.128***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Pro. Sociotropic	-0.006	-0.006	-0.005	-0.007	-0.007	-0.007
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
State Efficacy	0.161***	0.161***	0.161***	0.162***	0.161***	0.161***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Personal Efficacy	0.47***	0.47***	0.47***	0.469***	0.472***	0.471***
	(0.063)	(0.063)	(0.064)	(0.063)	(0.064)	(0.064)
Party ID	0.025**	0.025**	0.025**	0.025**	0.025**	0.025**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.014	0.014	0.014	0.014	0.014	0.014
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Age	-0.012	-0.012	-0.012	-0.012	-0.012	-0.012
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.071*	0.071*	0.071*	0.071*	0.071*	0.071*
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Income	0.1***	0.1***	0.1***	0.1***	0.1***	0.1***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Education	0.169***	0.169***	0.169***	0.169***	0.169***	0.169***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Knowledge	0.22***	0.216***	0.22***	0.372***	0.22***	0.161
	(0.028)	(0.048)	(0.028)	(0.110)	(0.028)	(0.204)
Economy Salient	0.017	0.017	0.018	0.019	0.017	0.017
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Inpartisan	0.137***	0.137***	0.138***	0.14***	0.138***	0.138***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Macro-level						
Inflation	0.015	0.013				
	(0.017)	(0.024)				
Inflation x Knowledge		0.001				
		(0.009)				
Unemployment			0.023	0.064‡		
			(0.041)	(0.048)		
Unemployment x Knowledge				-0.025†		
				(0.016)		
Consumer sentiment					-0.006‡	-0.007
					(0.005)	(0.006)
Consumer sent. x Knowledge						0.001
						(0.002)
Security Environment	-0.138	-0.14	0.151	0.18	0.338	0.34
	(0.435)	(0.435)	(0.453)	(0.425)	(0.432)	(0.431)
Random effects						
Intercept	(0.127)	(0.127)	(0.065)	(0.054)	(0.045)	(0.045)
Personal Efficacy	(0.172)	(0.172)	(0.176)	(0.172)	(0.177)	(0.177)
σ^2	0.995	0.995	0.995	0.995	0.995	0.995
ρ	0.022	0.022	0.008	0.006	0.003	0.003

N=22078, Years = 12. Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

Table 10: Contextual security determinants: main and nonlinear effects (I)

	MIDS	MIDS	Hostility Levels	Hostility Levels
	(1)	(2)	(3)	(4)
Micro-level				
Intercept	-1.493***	-1.028‡	-1.555***	-0.621
	(0.338)	(0.840)	(0.386)	(1.287)
Retro. Sociotropic	-0.128***	-0.128***	-0.125***	-0.126***
	(0.032)	(0.032)	(0.032)	(0.032)
Pro. Sociotropic	-0.006	-0.006	-0.005	-0.005
	(0.032)	(0.032)	(0.031)	(0.031)
State Efficacy	0.158***	0.158***	0.161***	0.161***
	(0.025)	(0.025)	(0.027)	(0.027)
Personal Efficacy	0.467***	0.469***	0.47***	0.473***
	(0.065)	(0.065)	(0.064)	(0.065)
Party ID	0.025**	0.025**	0.025**	0.025**
	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.013	0.013	0.014	0.014
	(0.017)	(0.017)	(0.017)	(0.017)
Age	-0.012	-0.012	-0.012	-0.012
	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.069*	0.069*	0.071*	0.071*
	(0.038)	(0.038)	(0.038)	(0.038)
Income	0.101***	0.101***	0.1***	0.1***
	(0.020)	(0.020)	(0.019)	(0.019)
Education	0.168***	0.168***	0.169***	0.169***
	(0.015)	(0.015)	(0.015)	(0.015)
Knowledge	0.22***	0.22***	0.22***	0.22***
	(0.029)	(0.029)	(0.028)	(0.028)
Economy Salient	0.023	0.022	0.018	0.017
	(0.042)	(0.043)	(0.045)	(0.045)
Inpartisan	0.139***	0.139***	0.138***	0.137***
	(0.038)	(0.038)	(0.038)	(0.038)
Macro-level				
MIDS	-0.004	-0.068		
	(0.013)	(0.099)		
$MIDS^2$		0.002		
		(0.003)		
Hostility Levels			0.000	-0.042
			(0.006)	(0.055)
Hostility $Levels^2$				0.000
				(0.001)
Unemployment	0.023	0.017	0.02	0.012
	(0.039)	(0.042)	(0.042)	(0.044)
Random effects				
(Intercept)	(0.107)	(0.087)	(0.087)	(0.087)
95% CI on Intercept	[0.154,0.217]	[0.232,0.298]	[0.151,0.2]	[0.312,0.389]
Personal Efficacy	(0.181)	(0.181)	(0.176)	(0.175)
95% CI on Efficacy	[0.528,0.694]	[0.529,0.695]	[0.531,0.693]	[0.533,0.693]
σ^2	0.995	0.995	0.995	0.995
ρ	0.014	0.01	0.011	0.011
N	22078	22078	22078	22078
Years	12	12	12	12

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

Table 11: Contextual security determinants: main and nonlinear effects (II)

	Casualties (5)	Casualties (6)	Failures (7)	Failures (8)
Micro-level				
Intercept	-1.646***	-1.556***	-1.411***	-1.245**
	(0.243)	(0.281)	(0.300)	(0.473)
Retro. Sociotropic	-0.129***	-0.129***	-0.127***	-0.126***
	(0.032)	(0.032)	(0.032)	(0.032)
Pro. Sociotropic	-0.007	-0.008	-0.005	-0.005
	(0.032)	(0.032)	(0.031)	(0.031)
State Efficacy	0.158***	0.158***	0.161***	0.161***
	(0.025)	(0.025)	(0.027)	(0.027)
Personal Efficacy	0.47***	0.468***	0.471***	0.473***
	(0.064)	(0.065)	(0.064)	(0.065)
Party ID	0.025**	0.025**	0.025**	0.025**
	(0.009)	(0.009)	(0.009)	(0.009)
Ideology	0.013	0.013	0.014	0.014
	(0.017)	(0.017)	(0.017)	(0.017)
Age	-0.012	-0.012	-0.012	-0.012
	(0.012)	(0.012)	(0.012)	(0.012)
Male	0.069*	0.069*	0.071*	0.071*
	(0.038)	(0.039)	(0.038)	(0.038)
Income	0.101***	0.101***	0.1***	0.1***
	(0.020)	(0.020)	(0.019)	(0.019)
Education	0.168***	0.168***	0.169***	0.169***
	(0.015)	(0.015)	(0.015)	(0.015)
Knowledge	0.22***	0.22***	0.219***	0.219***
	(0.029)	(0.029)	(0.028)	(0.028)
Economy Salient	0.022	0.022	0.018	0.018
	(0.043)	(0.043)	(0.045)	(0.045)
Inpartisan	0.14***	0.139***	0.138***	0.138***
	(0.038)	(0.038)	(0.038)	(0.038)
Macro-level				
Casualties	0.031	-0.051		
	(0.031)	(0.142)		
Casualties^2		0.016		
		-0.027		
Failures to Win			-0.011	-0.045
			(0.013)	(0.067)
Failures^2				0.002
				-0.003
Unemployment	0.031	0.027	0.016	0.013
	(0.036)	(0.036)	(0.039)	(0.042)
Random effects				
(Intercept)	(0.072)	(0.062)	(0.117)	(0.099)
95% CI on Intercept	[0.143,0.182]	[0.158,0.192]	[0.162,0.235]	[0.192,0.259]
Personal Efficacy	(0.182)	(0.183)	(0.175)	(0.175)
95% CI on Efficacy	[0.528,0.695]	[0.527,0.696]	[0.532,0.693]	[0.532,0.693]
σ^2	0.995	0.995	0.995	0.995
ρ	0.007	0.005	0.018	0.013
N	22078	22078	22078	22078
Years	12	12	12	12

Micro-level: *** $p < 0.001$ Macro-level: $\dagger p < 0.10$

Table 12: Contextual security determinants: other economic controls (I)

	Inflation (1)	Inflation (2)	Inflation (3)
Micro-level			
(Intercept)	-1.632***	-1.735***	-1.512***
	(0.313)	(0.391)	(0.255)
Retro. Sociotropic	-0.128***	-0.127***	-0.129***
	(0.032)	(0.032)	(0.031)
Pro. Sociotropic	(0.006)	(0.006)	(0.006)
	(0.031)	(0.031)	(0.031)
State Efficacy	0.161***	0.161***	0.161***
	(0.027)	(0.027)	(0.027)
Personal Efficacy	0.471***	0.471***	0.472***
	(0.062)	(0.062)	(0.061)
Party ID	0.025	0.025	0.025
	(0.009)	(0.009)	(0.009)
Ideology	0.014	0.014	0.014
	(0.017)	(0.017)	(0.017)
Age	(0.012)	(0.012)	(0.012)
	(0.012)	(0.012)	(0.012)
Male	0.071	0.071	0.071
	(0.038)	(0.038)	(0.038)
Income	0.1***	0.1***	0.1***
	(0.019)	(0.019)	(0.019)
Education	0.169***	0.169***	0.169***
	(0.015)	(0.015)	(0.015)
Knowledge	0.341***	0.362	0.321***
	(0.106)	(0.127)	(0.080)
Economy Salient	0.017	0.017	0.016
	(0.045)	(0.045)	(0.045)
Inpartisan	0.137***	0.137***	0.138***
	(0.038)	(0.038)	(0.038)
Macro-level			
MIDS	0.010		
	(0.019)		
MIDS x Knowledge	-0.008‡		
	(0.007)		
Hostility levels		0.006	
		(0.008)	
Hostility x Knowledge		(0.003)	
		(0.003)	
Failures to Win			0.001
			(0.018)
Failures x Know			(0.009)
			(0.007)
Inflation	0.015	0.015	0.017
	(0.017)	(0.018)	(0.016)
Random effects			
(Intercept)	(0.121)	(0.108)	(0.143)
95% CI on Intercept	[0.134,0.199]	[0.125,0.179]	[0.143,0.226]
Personal Efficacy	(0.167)	(0.166)	(0.163)
95% CI on Efficacy	[0.536,0.69]	[0.536,0.689]	[0.538,0.688]
σ^2	0.995	0.995	0.995
ρ	0.019	0.016	0.025
N	22078	22078	22078
Years	12	12	12

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

Table 13: Contextual security determinants: other economic controls (II)

	ICS (4)	ICS (5)	ICS (6)
Micro-level			
(Intercept)	-1.110	-1.145	-1.165
	(0.459)	(0.493)	(0.474)
Retro. Sociotropic	-0.132***	-0.131***	-0.130***
	(0.032)	(0.032)	(0.032)
Pro. Sociotropic	(0.007)	(0.008)	(0.006)
	(0.032)	(0.032)	(0.031)
State Efficacy	0.157***	0.157***	0.160***
	(0.025)	(0.025)	(0.027)
Personal Efficacy	0.466***	0.465***	0.472***
	(0.064)	(0.064)	(0.062)
Party ID	0.025	0.025	0.025
	(0.009)	(0.009)	(0.009)
Ideology	0.013	0.013	0.014
	(0.017)	(0.017)	(0.017)
Age	(0.012)	(0.012)	(0.012)
	(0.012)	(0.012)	(0.012)
Male	0.069	0.069	0.071
	(0.038)	(0.038)	(0.038)
Income	0.101***	0.101***	0.100***
	(0.019)	(0.019)	(0.019)
Education	0.168***	0.168***	0.169***
	(0.015)	(0.015)	(0.015)
Knowledge	0.348***	0.368	0.321***
	(0.105)	(0.127)	(0.080)
Economy Salient	0.022	0.022	0.017
	(0.042)	(0.042)	(0.045)
Inpartisan	0.139***	0.139***	0.138***
	(0.038)	(0.038)	(0.038)
Macro-level			
MIDS	0.024		
	(0.020)		
MIDS x Knowledge	-0.009		
	(0.007)		
Hostility levels		0.013†	
		(0.009)	
Hostility x Knowledge		(0.004)	
		(0.003)	
Failures to Win			0.012
			(0.020)
Failures x Know			-0.009
			(0.007)
Consumer Sentiment	-0.007†	-0.009†	(0.004)
	(0.005)	(0.005)	(0.005)
Random effects			
(Intercept)	(0.071)	(0.056)	(0.082)
95% CI on Intercept	[0.223,0.275]	[0.222,0.262]	[0.21,0.268]
Personal Efficacy	(0.178)	(0.180)	(0.168)
95% CI on Efficacy	[0.529,0.693]	[0.528,0.694]	[0.536,0.69]
σ^2	0.995	0.995	0.995
ρ	0.007	0.005	0.010
N	22078	22078	22078
Years	12	12	12

Micro-level: *** $p < 0.001$ Macro-level: †$p < 0.10$

References

Anderson, Christopher J. and Yuliya V. Tverdova. 2003. "Corruption, Political Allegiances, and Attitudes Toward Government in Contemporary Democracies." *American Journal of Political Science* 47(1):91–109.

Bartels, Brandon L. 2009. "The Constraining Capacity of Legal Doctrine on the U.S. Supreme Court." *American Political Science Review* 103(3):474–495.

Beck, Nathaniel. 2001. "Time-Series-Cross-Section Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4:271–93.

Bennett, D. Scott and Allan C Stam. 1998. "The Declining Advantages of Democracy: A Combined Model of War Outcomes and Duration." *Journal of Conflict Resolution* 42(3):344–366.

Braumoeller, Bear F. 2006. "Explaining Variance: Or, Stuck in a Moment We Can't Get Out Of." *Political Analysis* 14(3):268–290.

Falleti, Tulia G. and Julia F. Lynch. 2009. "Context and Causal Mechanisms in Political Analysis." *Comparative Political Studies* 42(9):1143–1166.

Fordham, Benjamin O. 2008. "Economic Interests and Public Support for American Global Activism." *International Organization* 62(1):163–182.

Gelman, Andrew, Boris Shor, Joseph Bafumi and David Park. 2007. "Rich State, Poor State, Red State, Blue State: What's the Matter with Connecticut?" *Quarterly Journal of Political Science* 2(4):345–367.

Gronke, Paul and John Brehm. 2002. "History, Heterogeneity, and Presidential Approval: A Modified ARCH Approach." *Electoral Studies* 21(3):425–452.

Hetherington, Marc J. and Michael Nelson. 2003. "Anatomy of a Rally Effect: George W. Bush and the War on Terrorism." *PS: Political Science & Politics* 36(1):37–42.

Hill, Peter W. and Kenneth J. Rowe. 1996. "Multilevel Modelling in School Effectiveness Research."

School Effectiveness and School Improvement 7(1):1–34.

Holsti, Ole R. 1979. "The Three-Headed Eagle: The United States and System Change." *International Studies Quarterly* 23(3):339–359.

Honaker, James, Gary King and Matthew Blackwell. 2009. "Amelia II: A Program for Missing Data." http://gking.harvard.edu/amelia.

Hurwitz, Jon and Mark Peffley. 1987. "How Are Foreign Policy Attitudes Structured?" *American Political Science Review* 81(4):1099–1120.

Hutchison, Marc L. and Douglas M. Gibler. 2007. "Political Tolerance and Territorial Threat: A Cross-National Study." *Journal of Politics* 69(1):128–142.

Kam, Cindy D. and Jennifer M. Ramos. 2008. "Joining and Leaving the Rally: Understanding the Surge and Decline in Presidential Approval Following 9/11." *Public Opinion Quarterly* 72(4):619–650.

Kedar, Orit and W. Phillips Shively. 2005. "Introduction to the Special Issue." *Political Analysis* 13(4):297–300.

Keele, Luke and Jennifer Wolak. 2008. "Contextual Sources of Ambivalence." *Political Psychology* 29(5):653–673.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69.

Lebo, Matthew J. and Janet M. Box-Steffensmeier. 2008. "Dynamic Conditional Correlations in Political Science." *American Journal of Political Science* 52(3):688–704.

Liberman, Peter. 2006. "An Eye for an Eye: Public Support for War Against Evildoers." *International Organization* 60(3):687–722.

Mueller, John E. 1973. *War, Presidents, and Public Opinion*. New York: John Wiley.

Newman, Brian and Andrew Forcehimes. 2010. ""Rally Round the Flag" Events for Presidential

Approval Research." *Electoral Studies* 29(1):144–154.

Parker, Suzanne L. 1995. "Toward an Understanding of "Rally" Effects: Public Opinion in the Persian Gulf War." *Public Opinion Quarterly* 59(4):526–546.

Rathbun, Brian C. 2007. "Hierarchy and Community at Home and Abroad: Evidence of a Common Structure of Domestic and Foreign Policy Beliefs in American Elites." *Journal of Conflict Resolution* 51(3):379–407.

Reiter, Dan and Allan C Stam. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.

Rubin, Donald. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Steenbergen, Marco R. and Bradford S. Jones. 2002. "Modeling Multilevel Data Structures." *American Journal of Political Science* 46(1):218–237.

Wittkopf, Eugene R. 1986. "On the Foreign Policy Beliefs of the American People: A Critique and Some Evidence." *International Studies Quarterly* 30(4):425–445.

Wittkopf, Eugene R. 1990. *Faces of internationalism: public opinion and American foreign policy*. Duke University Press.