**SUPPLEMENTARY APPENDIX**

**Coding Scheme**

Content Analysis of NYT Articles: We hired a team of research assistants to perform the conent analysis of the NYT for the years spanning1980-2002. The remaining time period (2000-2010), was coded using an automated content analysis. The automated content analysis used machine learning techniques, using the text classification package Rtexttools (Jurka et al 2012) and incorporated information from the hand-coded articles classification for the articles before 2000. Out of the nine algorithms available in the package, our research assistant (RA) used the following six techniques to classify each of the variables in the database: maximum entropy (Jurka 2012), support vector machine (Dimitriadou et al 2011), glmnet (Friedman, Hastie, and Tibshirani 2010), random forest (Liaw and Weiner 2002), boosting (Tuszynski, 2011), and classification trees (Ripley 2011). For every variable, the following procedure was used: First, all the articles in the database were transformed into a unique matrix of dimensions *n* x *m*, where *n* denotes the number of articles and *m* is the count of different words in the database.

To reduce the computational time and space, the procedure eliminated those columns in the matrix for words that appear in less than five percent of the articles. Second, in order to train each of the algorithms for text classification, the training set was defined as a subset of the database that only considers the hand-coded classification for those articles between 1980 and 1990. In this stage, each technique processed the probability that a word belongs to a given class of a variable given the content of the articles and their assigned category by the research assistants. Third, it uses the resultant syntax from the training set to classify the articles from 2000 to 2010. That is, each algorithm uses the information from the training set to make individual predictions about the unclassified articles. Finally, to assign a category out of the predictions of each technique, our research assistant used an ensemble agreement to assign each category, which refers to whether multiple algorithms coincide in the class of an event.

Regarding the classification of the immigrant's country of origin in the article, our RA considered a list of the top 30 immigrant groups in the US. The list is below. From each of the groups, our RA created a list of words that are commonly used to identify the origin of each group. For example, to identify whether the immigrants in the article come from Ukraine, the following list of words was created: Ukraine, Ukranians, Ukranian-American, and Ukranian-born. If any of these words appear in a given article, it was classified as mentioning Ukranian immigrants. The classification of every article was not exclusive and detected more than one nationality mentioned in the article.

We also performed a test of intercoder reliability between the automated dataset and the hand-coded dataset. This test of intercoder reliability, calculated in the Rtexttools (Jurka et al 2012) package, indicates that the two datasets coincide 70-75% of the time. This high degree of agreement makes us confident that there is nothing substantively different from these two datasets.

Framing Code:

**Tone Frames:** tone of content towards immigration:

1. Negative 2. Neutral 3. Positive

**Issue Frames:** general content regarding immigration:

1. Economic Issue (jobs, unemployment)
2. Policy Issue
3. Health Issue (of the immigrant)
4. Family reunification
5. National (Homeland) Security Issue
6. Crime (e.g. arrests, human smuggling, drug trafficking, border arrests or skirmishes, illegal aspects associated with in general but that is not terrorism or homeland security).
7. Social Welfare (public services, welfare benefits)
8. Social fabric, culture of the U.S.

**Table A1: The Effect of Immigration Frames on the White Macropartisanship- Assessing Issue Content**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Percent Democratic Identifiers** | **Percent Independents** | **Percent Republican Leaners** |
| ***Immigration Frames*** |  |  |  |
| *TONE* |  |  |  |
| Negative Tone | -0.56 (.62) | 0.81 (.71) | -.18 (1.02) |
| *GROUP IMAGE* |  |  |  |
| Latino | -4.08 (1.63)\* | 4.83 (1.91)\* | 7.27 (3.02)\* |
| *ISSUE CONTENT* |  |  |  |
| Crime | -1.14 (1.88) | 2.39 (2.08) | -.76 (2.96) |
| Economy | 2.13 (3.20) | -5.32 (3.22) | -.22 (4.54) |
| Immigration Policy | 2.67 (1.01)\*\* | -1.05 (.92) | 1.91 (1.32) |
| ***Agenda Setting*** |  |  |  |
| Volume of Coverage | .001 (.006) | -.005 (.006) | .01 (.01) |
| ***Other Controls*** |  |  |  |
| Presidential Approval | -.05 (.02)\* | -.01 (.01) | .06 (.02)\* |
| Unemployment Rate | 0.34 (.28) | .52 (.15)\*\* | -.36 (.27) |
| Constant | 31.0 (2.12)\*\* | 28.4 (1.3)\*\* | 43.9 (2.29)\*\* |
| N | 115 | 115 | 94 |
| R2 | 0.81 | 0.54 | 0.82 |

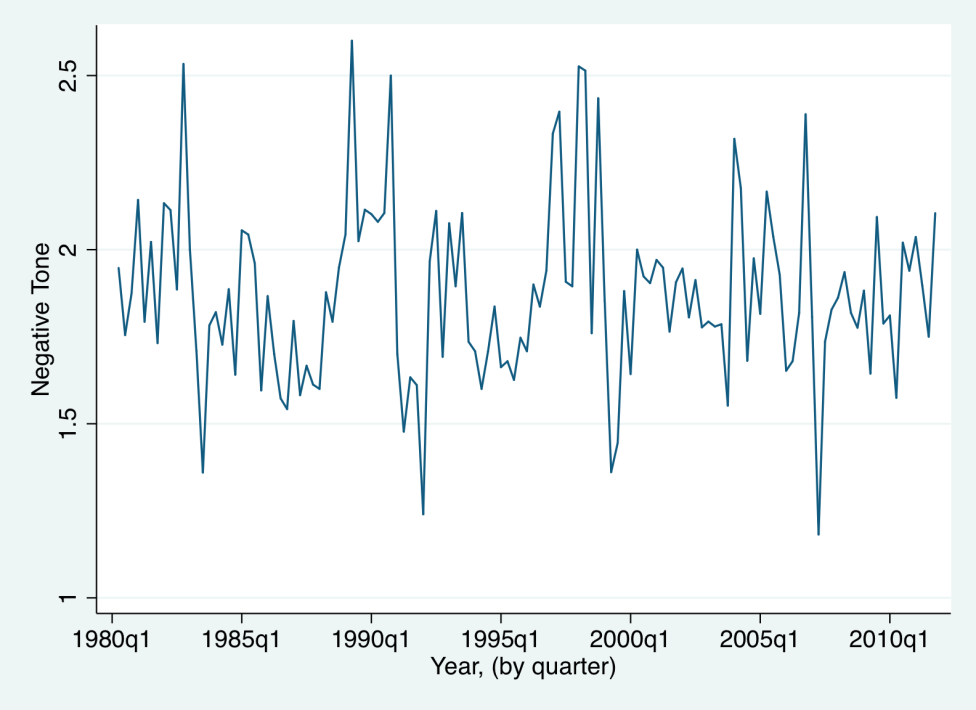
\*\*p<.01, \*p<.05

Coefficients are Prais-Winsten AR(1) regression estimates. Standard errors in parentheses

**Figure A1. Volume of NYT Coverage of Immigration, by Year and Quarter**



**Figure A2: Balance of Negative vs Positive Frames, by Year and Quarter**

****

**Controlling for Other Policy Issues**

By restricting our attention to immigration and ignoring media coverage of other salient issues, we may be unfairly biasing our results in favor of significant findings. To address this concern, we collected and incorporated data on NYT coverage of welfare, terrorism, and war - three issues that received widespread coverage over this period of time and three issues that many might view as being primarily responsible for Republican gains over the same period.

Our coding scheme was similar but due to time constraints we machine coded all articles for these three other issues using Liu’s opinion lexicon (Liu 2015). As before, we looked at both the tone of the articles (the average balance of negative vs positive words) and the total amount of coverage devoted to each issue each quarter. Table A2 displays the results of the analysis.

**Table A2: The Effect of Immigration Frames on the White Macropartisanship- CONTROLLING FOR MEDIA COVERAGE IN OTHER POLICY AREAS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Democrats Relative to Republicans** | **Percent Democratic Identifiers** | **Percent Independents** | **Percent Republican Leaners** |
| ***IMMIGRATION*** |  |  |  |  | |
| Latino Group Image | -14.4 (4.87)\*\* | -4.50 (1.76)\* | 6.02 (1.94)\*\* | 8.60 (3.21)\*\* | |
| Tone of Coverage | -1.52 (1.68) | -.849 (.661) | .719 (.747) | .336 (1.12) | |
| Crime Focus | -1.12 (4.87) | -1.14 (2.06) | 1.81 (2.29) | -.126 (3.27) | |
| Volume of Coverage | .014 (.023) | .003 (.006) | -.004 (.006) | -.013 (.015) | |
| **TERRORISM** |  |  |  |  | |
| Tone | -49.1 (45.5) | -32.2 (18.7) | 11.2 (20.6) | 19.3 (30.0) | |
| Volume of Coverage | -.000 (.000) | -.011 (1.19) | .000 (.001) | .000 (.001) | |
| Tone\*Volume | .000 (.001) | -.000 (.001) | .000 (.000) | .000 (.001) | |
| ***WAR*** |  |  |  |  | |
| Tone | -163 (250) | -20.19 (105.9) | 43.3 (120) | 80.9 (167) | |
| Volume of Coverage | -.000 (.001) | -.085 (.365) | .002 (.004) | .000 (.001) | |
| Tone\*Volume | -.000 (.001) | -.000 (.004) | -.000 (.000) | -.000 (.001) | |
| ***WELFARE*** |  |  |  |  | |
| Tone | 152 (107) | 4.03 (32.6) | -11.2 (36.7) | -96.7 (70.1) | |
| Volume of Coverage | .000 (.000) | .000 (.000) | -.002 (.001) | .000 (.002) | |
| Tone\*Volume | -.002 (.003) | .000 (.001) | -.002 (.001) | .001 (.002) | |
| Other Controls |  |  |  |  | |
| Presidential Approval | -.106 (.043)\* | -.061 (.021)\*\* | .007 (.017) | .061 (.026)\* | |
| Unemployment Rate | .735 (.521) | .518 (.284) | .281 (.182) | -.276 (.298) | |
| Constant | -11.7 (5.51)\* | 30.11 (2.57)\*\* | 30.3 (2.28)\*\* | 42.4 (3.47)\*\* | |
| N | 94 | 115 | 115 | 94 | |
| R2 | .23 | .80 | .61 | .81 | |

\*\*p<.01, \*p<.05

Coefficients are Prais-Winsten AR(1) regression estimates. Standard errors in parentheses.

As the estimates from Table A2 demonstrate, the incorporation of other salient issues does not alter our main findings. Immigration coverage continues to help explain shifts in aggregate partisanship.

There are also some weak signs that more positive coverage of terrorism increases the percentage of white Americans who identify with the President’s Party but the effects are not at all robust across the different dependent variables.

In addition, to make sure that the effects of immigration coverage and lack of effects of coverage of other issues on white macropartisanship were not the result of a difference in coding, we also machine coded all the immigration articles using Liu’s opinion lexicon. We then ran models identical to Table A2 that used the machine coded variable for immigration alongside the other issues. These models showed the exact same results with the same levels of significance.

**Error Correction Model**

In line with DeBoef and Keele (2008), we sought to reanalyze our data with a different form of time-series model to help confirm the basic pattern of findings. Specifically, we repeated the analysis with an error correction model (ECM). The results are displayed below in Table A3. Here we include a reduced set of significant estimators. Including the full range of other issue positions does not change our core findings.

**Table A3: The Effect of Immigration Frames on the White Macropartisanship- Assessing Issue Content- Using the ERROR CORRECTION MODEL**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Democrats Relative to Republicans** | **Percent Democratic Identifiers** | **Percent Independents** | **Percent Republican Leaners** |
| ***Latino Group Image*** |  |  |  |  | |
| Lag Latino | -16.8 (7.52)\* | -5.70 (1.93)\*\* | 1.96 (2.04) | 9.30 (4.72)\* | |
| Delta Latino | -14.3 (5.59)\* | -5.12 (1.66)\*\* | 5.38 (1.95)\*\* | 7.82 (3.60)\* | |
| **Presidential Approval** |  |  |  |  | |
| Lag Approval | -.15 (.03)\*\* | -.05 (.01)\*\* | .000 (.014) | .094 (.022)\*\* | |
| Delta Approval | -.06 (.04) | -.05 (.02)\*\* | .006 (.024) | .006 (.028) | |
| ***Tone of Terrorism*** |  |  |  |  | |
| Lag Tone Terrorism | -76.9 (18.6)\*\* | -26.9 (7.17)\*\* | -.642 (7.21) | 37.9 (11.3)\*\* | |
| Delta Tone Terrorism | -20.1 (27.8) | -15.6 (11.2) | 25.7 (13.3)\* | 1.90 (17.9) | |
| ***Past Partisanship*** |  |  |  |  | |
| Partisanship Lagged | -.46 (.08)\*\* | -.327 (.067)\*\* | -.571 (.089)\*\* | -.58 (.09)\*\* | |
| Constant | -11.7 (5.51)\* | 30.11 (2.57)\*\* | 17.4 (2.95)\*\* | 20.1 (3.24)\*\* | |
| N | 85 | 110 | 110 | 85 | |
| R2 | .31 | .25 | .30 | .35 | |

\*\*p<.01, \*p<.05 Standard errors in parentheses

As can been seen in Table A3, the ECM analysis confirms our core conclusion. A greater focus by the media on Latinos once again was related to fewer ties to the Democratic Party and greater attachment to the Republican Party. The ECM model is also helpful in that it can help us say something about the short- and long-run effects of immigration coverage. The ECM model indicates that immigration media coverage has both a significant temporary effect and a significant long-run impact. Specifically, the ECM model estimates that a ten percent increase in NYT coverage of Latinos is associated with an immediate 1.6 point shift in the balance of Democrats and Republicans in the nation and a 1.4 point long term shift in the partisan balance (model in first column). Calculating the long run multiplier, the ECM estimates that a ten percent increase in NYT coverage of Latinos is associated with a total 3.6 shift in the partisan balance over all future periods.

**Determining the Correct Lag**

To help justify our main Prais-Winsten AR(1) regression model and to help further our understanding of the temporal effects of immigration coverage, we estimated Schwarz’s Bayesian information criterion (SBIC) and Hannan and Quinn information criterion (HQIC) to obtain lag-order selection statistics for a series of vector autoregressions. These two test statistics indicated that the macropartisanship data has AR(1) lags (see tables in the supplementary online appendix). While there are various ways to determine order selection, we use the SBIC and HQIC statistics due to its theoretically desirable properties when compared to AIC and FPE (Lutkepohl 2005). As Lutkepohl (2005, 148–152) shows, “choosing to minimize the SBIC or the HQIC provides consistent estimates of the true lag order, p. In contrast, minimizing the AIC or the FPE will overestimate the true lag order with positive probability, even with an infinite sample size.” <http://www.stata.com/manuals13/tsvarsoc.pdf>

**DV: Democrat**

Selection-order criteria

Sample: 1981q1 - 2010q4, but with a gap Number of obs = 115

+---------------------------------------------------------------------+

lag LL LR df p FPE AIC HQIC SBIC

----+------------------------------------------------------------------

0 -264.818 5.95982 4.62292 4.63261 4.64679

1 -213.116 103.4\* 1 0.000 2.46767 3.74115 3.76053\* 3.78889\*

2 -211.839 2.5551 1 0.110 2.45581\* 3.73632\* 3.76539 3.80793

3 -211.118 1.4412 1 0.230 2.46781 3.74118 3.77993 3.83666

4 -210.822 .592 1 0.442 2.49828 3.75342 3.80187 3.87277

+----------------------------------------------------------------------

**DV: Independents**

Selection-order criteria

Number of obs = 108

+---------------------------------------------------------------------------+

lag LL LR df p FPE AIC HQIC SBIC

----+----------------------------------------------------------------------

0 -1004.37 5.38289 18.7105 18.7709 18.8595

1 -663.829 681.08 36 0.000 .019153 13.0709 13.4938\* 14.1139\*

2 -609.566 108.53 36 0.000 .013739\* 12.7327\* 13.5181 14.6698

3 -582.698 53.736 36 0.029 .016519 12.9018 14.0497 15.7329

4 -552.098 61.2\* 36 0.005 .018795 13.0018 14.5122 16.727

+---------------------------------------------------------------------------+

**DV: Republicans**

Selection-order criteria

Number of obs = 108

+---------------------------------------------------------------------------+

lag | LL LR df p FPE AIC HQIC SBIC |

----+----------------------------------------------------------------------|

0 | -1004.37 5.38289 18.7105 18.7709 18.8595 |

1 | -663.829 681.08 36 0.000 .019153 13.0709 13.4938\* 14.1139\* |

2 | -609.566 108.53 36 0.000 .013739\* 12.7327\* 13.5181 14.6698 |

3 | -582.698 53.736 36 0.029 .016519 12.9018 14.0497 15.7329 |

4 | -552.098 61.2\* 36 0.005 .018795 13.0018 14.5122 16.727 |

+---------------------------------------------------------------------------+

**References**

Liu, Bing. 2015. *Sentiment Analysis Mining Opinions, Sentiments, and Emotions.* New York: Cambridge University Press.

Lutkepohl, H. 1993. *Introduction to Multiple Time Series Analysis*. 2nd ed. New York: Springer.