The Gatekeeping Function: Data

Stuart Soroka
McGill University

This file is distributed alongside the three datasets used in the following paper:

Stuart Soroka. 2012. "The Gatekeeping Function: Distributions of Information in Media and the Real World," The Journal of Politics 74(2); 514-528.

TSdata.dta: The time series data used for Table 1.
NYT_unemp: Raw media data, *New York Times*, unemployment stories.
NYT_infl: Raw media data, *New York Times*, inflation stories.
NYT_intrate: Raw media data, *New York Times*, interest rate stories.
WPOST_unemp: Raw media data, *Washington Post*, unemployment stories.
WPOST_infl: Raw media data, *Washington Post*, inflation stories.

Note that media data files do not contain the raw text — just the date of articles, the number words, and net tone.

Table 1 can be reproduced using the following STATA syntax:

```
*unemployment recodes
gen dunemp=D.USunemp
gen dunemp_up=dunemp if dunemp>=0
recode dunemp_up  .=0 if dunemp~=.
gen dunemp_dn=dunemp if dunemp<0
recode dunemp_dn  .=0 if dunemp~=.

*SYMMETRIC MODEL
reg NYTtone_unemp L.dunemp L.NYTtone_unemp

*ASYMMETRIC MODEL
reg NYTtone_unemp L.dunemp_up L.dunemp_dn L.NYTtone_unemp
test L.dunemp_up=L.dunemp_dn
```

Details on the syntax used to generate distributions is available upon request.