

APPENDICES

to accompany

Extreme Weather Events and Rural-to-Urban Migration

Trudy Ann Cameron*

Department of Economics, University of Oregon

Raisa Saif

Program in Environmental Studies, University of Oregon

Eric Duquette

Economic Research Service, U.S. Department of Agriculture

Appendix A –

Can the censored migration information be recovered?

Suppose there were indications that the dependent variable in our model to explain migration flows could be used in “levels” form, rather than the log form implied by the standard gravity model. Then there would be an opportunity to make use of the reported aggregate flows to “other counties in the same state,” or to each of the four regions (net of the explicitly county-level flows). Any regression model that involved the levels of county-to-county flows on the left-hand side and a linear-in-parameters specification for the right-hand-side variables would permit the individual-destination county observations to be included as:

$$\begin{aligned} migration_{ijt} = & \log \left(g \left[X_{ijt} \right] \right) \\ & + \beta_1 \log(Pop_i) + \beta_2 \log(Pop_j) - 2\beta_3 \log(distance_{ij}) + \varepsilon_{ijt} \end{aligned} \quad (1)$$

Let J' be the number of counties over which an aggregated flow is calculated. Then these aggregated flows could likewise be included in the model by aggregating the explanatory variables in a similar fashion:

$$\begin{aligned} \sum_{j=1}^{J'} migration_{ijt} = & \sum_{j=1}^{J'} \log \left(g \left[X_{ijt} \right] \right) + \beta_1 (J') \log(Pop_i) \\ & + \beta_2 \sum_{j=1}^{J'} \log(Pop_j) - 2\beta_3 \sum_{j=1}^{J'} \log(distance_{ij}) + \sum_{j=1}^{J'} \varepsilon_{ijt} \end{aligned} \quad (2)$$

For these aggregated observations, assuming independent observations with i.i.d. errors, the variance of the error term would be $(J')^2$ times the error variance for all other county pairs.

Thus GLS estimates could be attained by scaling the aggregate migration flows and the

aggregated right-hand-side variables by J' , the number of counties over which the aggregate measures are calculated.

Unfortunately, for the uncensored sample of flows in our data, the maximized value of the log-likelihood for otherwise identical specifications that use $migration_{ijt}$ rather than $\log(migration_{ijt})$ as the dependent variable do not produce similar maximized values of the log-likelihood function. Rather than being off by just a few points, the log-likelihoods are on the order of -3,500,000 versus -500,000 for the “levels” versus “logs” specifications. It is not possible to argue that the “lin-log” specification is close enough to the “log-log” specification so that we can resort to the models in equations (5) and (6). Since the uncensored portion of the data indicates that we must use some transformation of the dependent variable, a conforming specification for the aggregated flows is not possible.

As a consequence, we will work with only the uncensored flows, so that our results pertain only to “significant flows.” We assess the influence of the smaller flows by exploring the data for parameter differences across an ordinary regression and an explicated truncated regression model.

Table A. – Correlations between types of extreme weather events, by measure of impact (obs=576680 migration flows)

Any events?	floods	drought	hail	heat	hurricane	sev.storm	tornadoes	wildfires	winds
origin_floods	1								
origin_drought	0.0281	1							
origin_hail	0.1675	0.0955	1						
origin_heat	0.061	0.0446	-0.0051	1					
origin_hurricane	0.0026	-0.0237	0.0129	-0.0309	1				
origin_sev.storm	-0.4263	-0.0788	-0.2483	-0.1506	-0.1375	1			
origin_tornadoes	0.1183	0.0541	0.1441	0.0158	0.1493	0.0055	1		
origin_wildfires	0.0996	0.0666	0.0033	0.0998	-0.0276	-0.0571	0.0902	1	
origin_winds	-0.2219	-0.041	-0.1293	-0.0784	-0.0716	-0.1342	-0.132	-0.0582	1
origin_winter	0.1835	0.0682	0.1314	0.136	-0.1003	-0.0893	-0.0019	0.0706	0.0185
# Events	floods	drought	hail	heat	hurricane	sev.storm	tornadoes	wildfires	winds
origin_floods	1								
origin_drought	-0.0182	1							
origin_hail	0.2458	0.0679	1						
origin_heat	0.0506	0.0469	0.0101	1					
origin_hurricane	-0.0417	-0.014	0.0068	-0.0336	1				
origin_sev.storm	-0.138	-0.0415	-0.0867	-0.0741	-0.0796	1			
origin_tornadoes	0.1584	0.0458	0.2121	0.0062	0.1844	0.0153	1		
origin_wildfires	0.3632	0.0164	-0.0052	0.1389	-0.0289	-0.0408	0.0905	1	
origin_winds	-0.0722	-0.0217	-0.0453	-0.0387	-0.0416	-0.0579	-0.0703	-0.0262	1
origin_winter	0.1989	0.0388	0.1245	0.0917	-0.0756	-0.0542	0.0042	0.0513	0.0149

Alternative measures for the difference in urbanization across counties

We have explored a number of candidate functions for $[UrbanizationDifference_{ij}]$ beyond the simple indicator variable, $1(PropUrban_j > PropUrban_i)$, discussed in the body of the paper. Other candidate measures include

- A continuous differential variable, $[PropUrban_j - PropUrban_i]$, that measures the extent to which the urbanization level in the destination county exceeds the urbanization level in the origin county. This variable conveys the degree to which urbanization differs, not just the sign of the difference, but forces the migration function to be linear in this difference. Likewise, this specification imposes the assumption that differences in destination urbanization and differences in origin urbanization have equal but opposite effects on expected migration between an origin-destination county pair.
- Ten discrete indicators for deciles of the distribution of $[PropUrban_j - PropUrban_i]$. This set of indicators allows the migration function to depend nonlinearly on the urbanization differential between the destination county and the origin county, but retains the assumption that the difference is what drives migration behavior.
- Separate sets of ten discrete indicators for deciles of the distributions of $PropUrban_i$ and the distribution of $PropUrban_j$. This specification is much more “non-parametric” and flexible, allowing different effects for 100 different combinations of origin and destination urbanization levels. This specification has the potential to reveal complex patterns, but the need to estimate 100 distinct coefficients affects the power of the individual estimates. With ten types of extreme weather events and seven candidate measures for the impact of each event, a fully general model using this option would

involve potentially $2 \times 7 \times 10 \times 100 = 14,000$ coefficients of the β_{4k} and β_{5k} variety. This parameter space is too large to be manageable, even with roughly five million observations.

Figures

Figure A.1, in ten parts, shows the variations over time in extreme weather events for 2006 through 2010 by county in the lower 48 states of the U.S. Note that occasional missing data for some counties is indicated by white spaces on these maps. (0 (grey)= no such event; 1 (other color)=at least one recorded event)

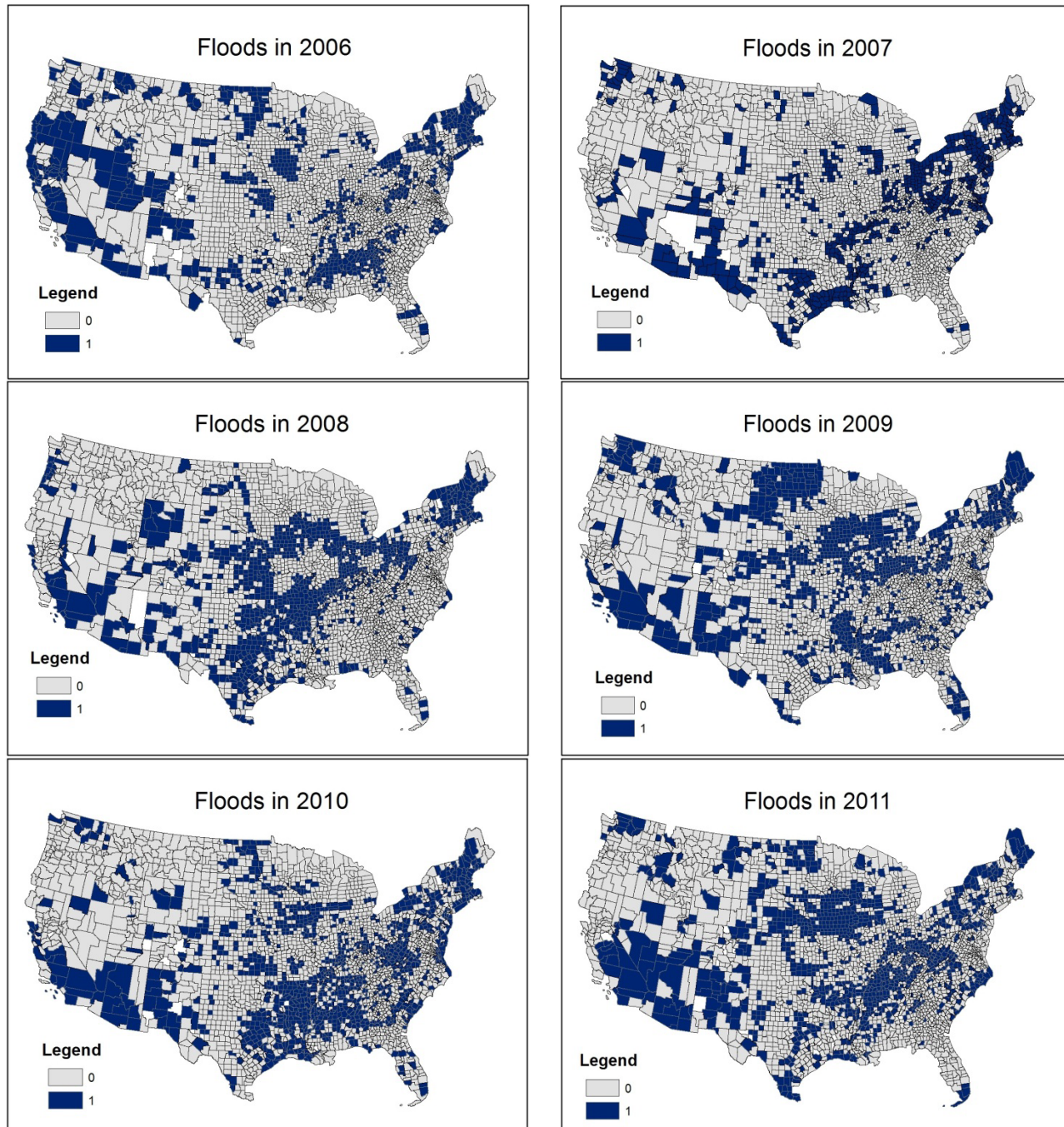


Figure A.1a

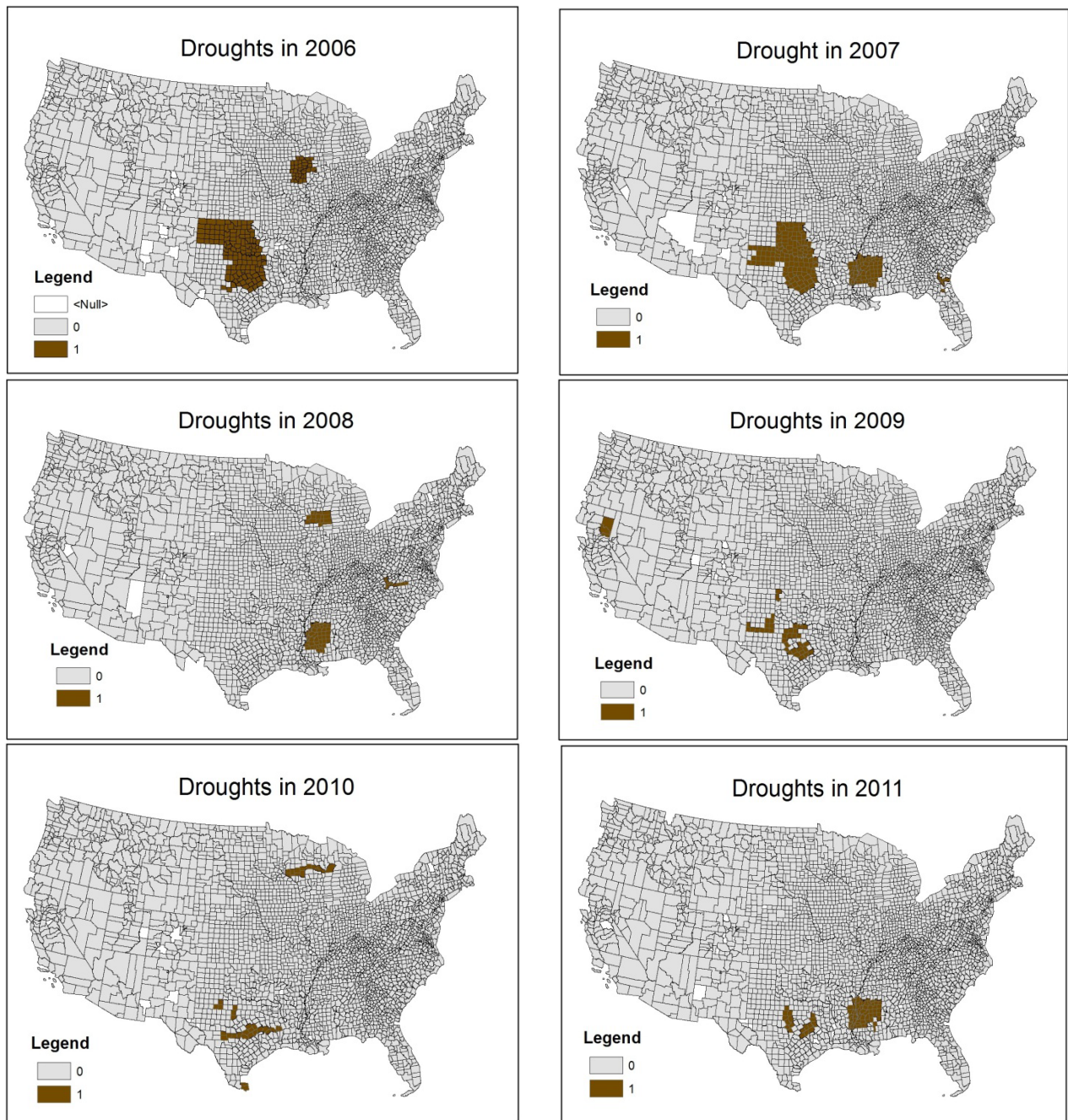


Figure A.1b

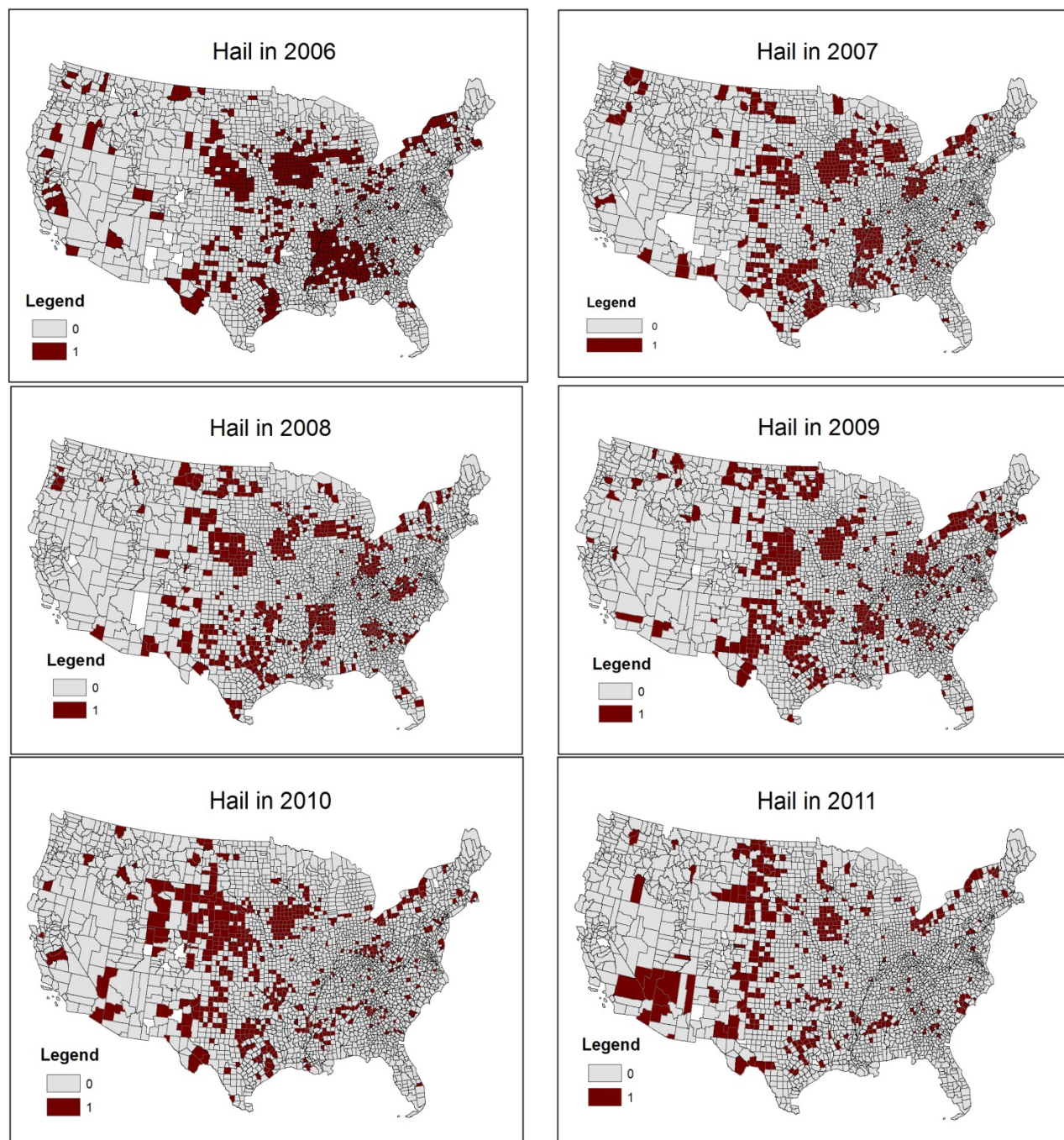


Figure A.1c

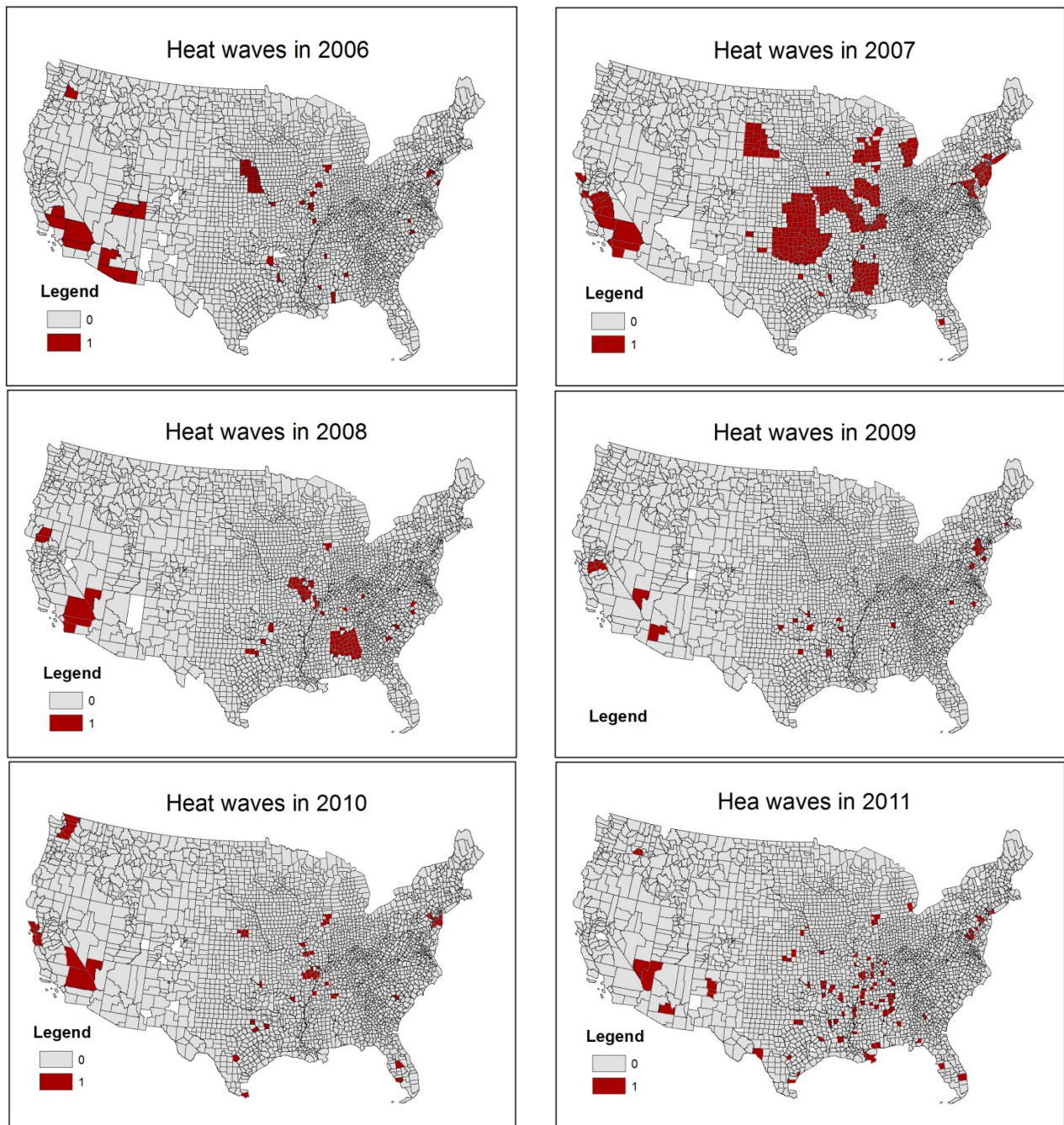


Figure A.1d

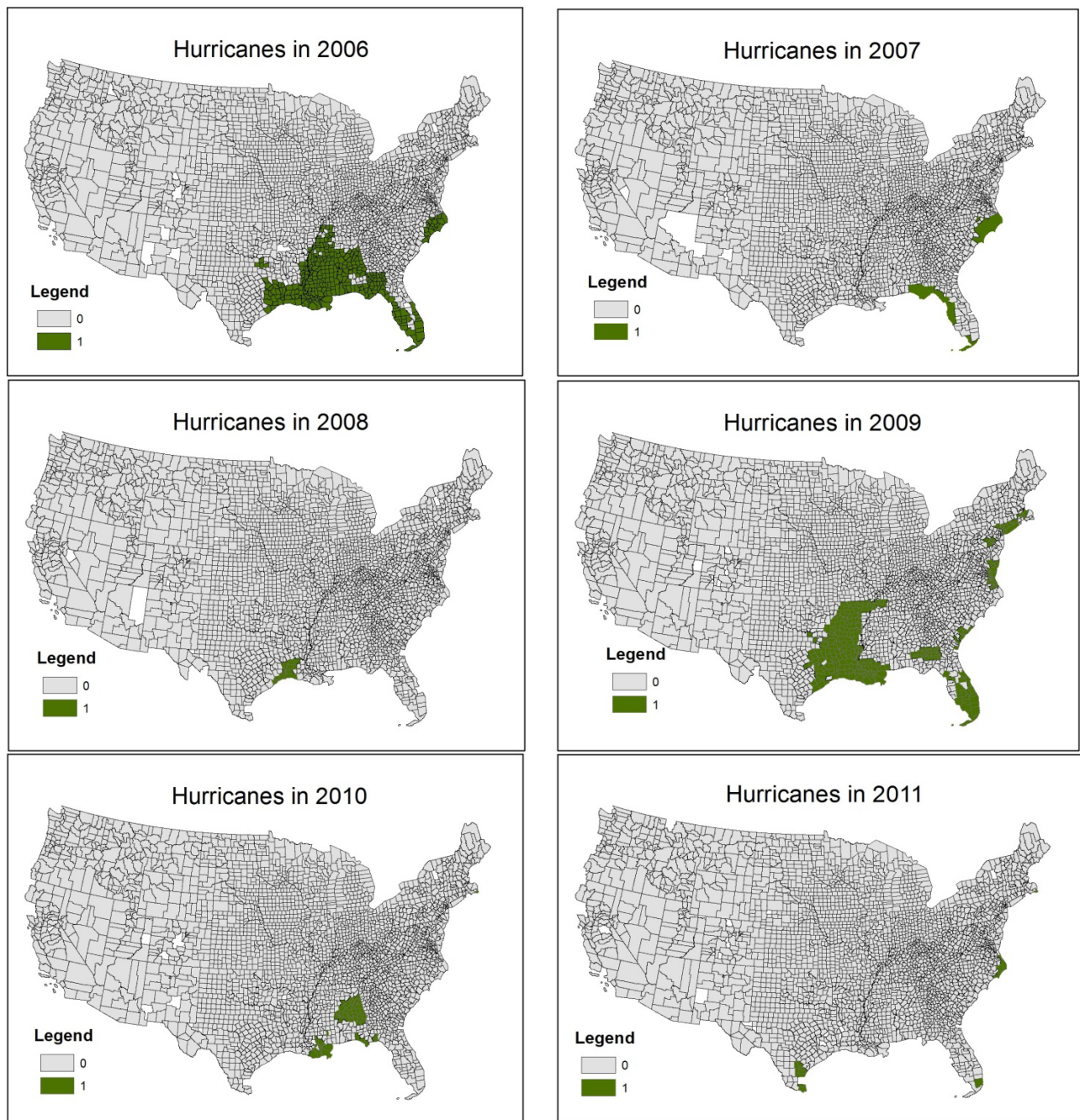


Figure A.1e

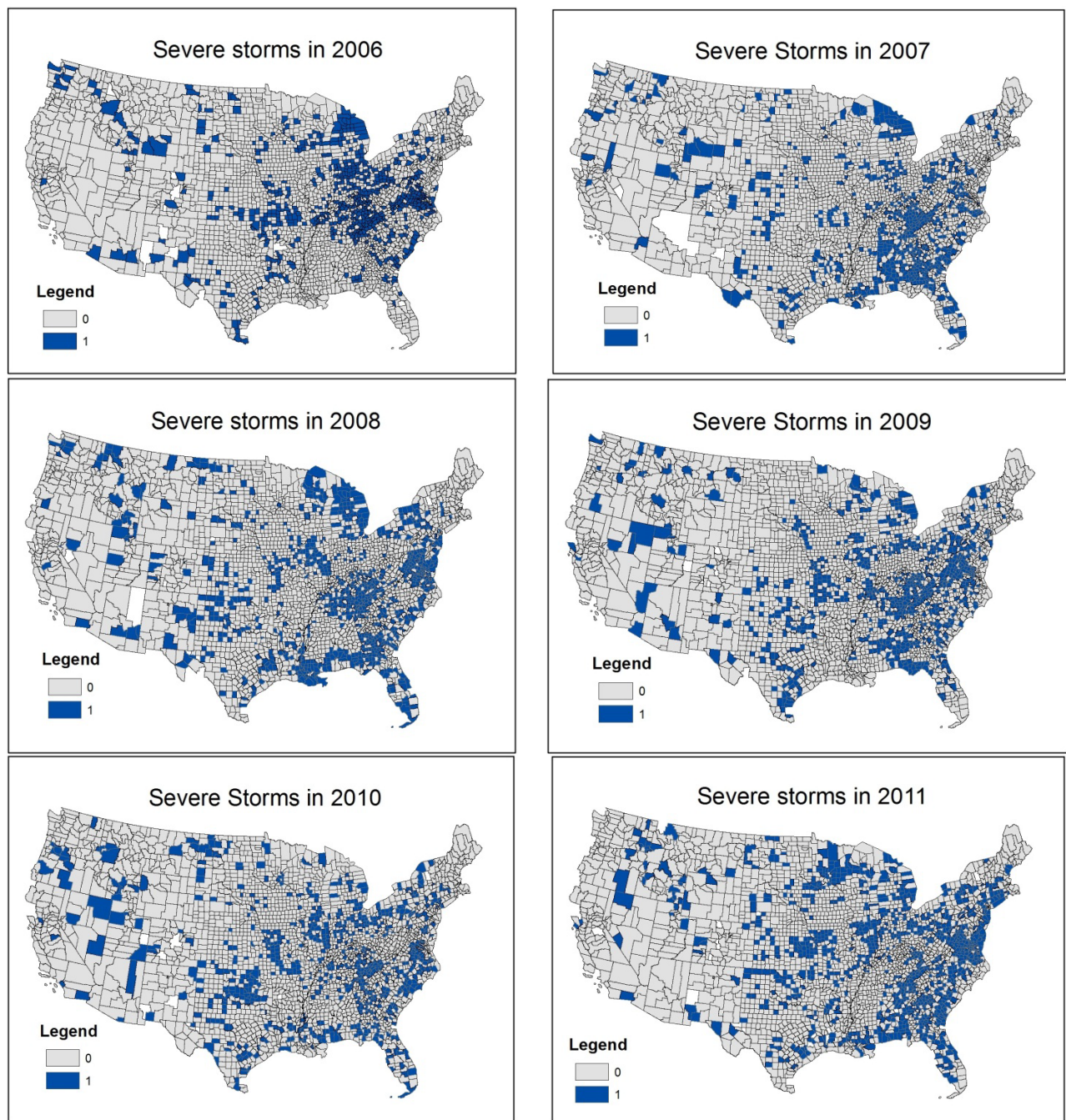


Figure A.1f

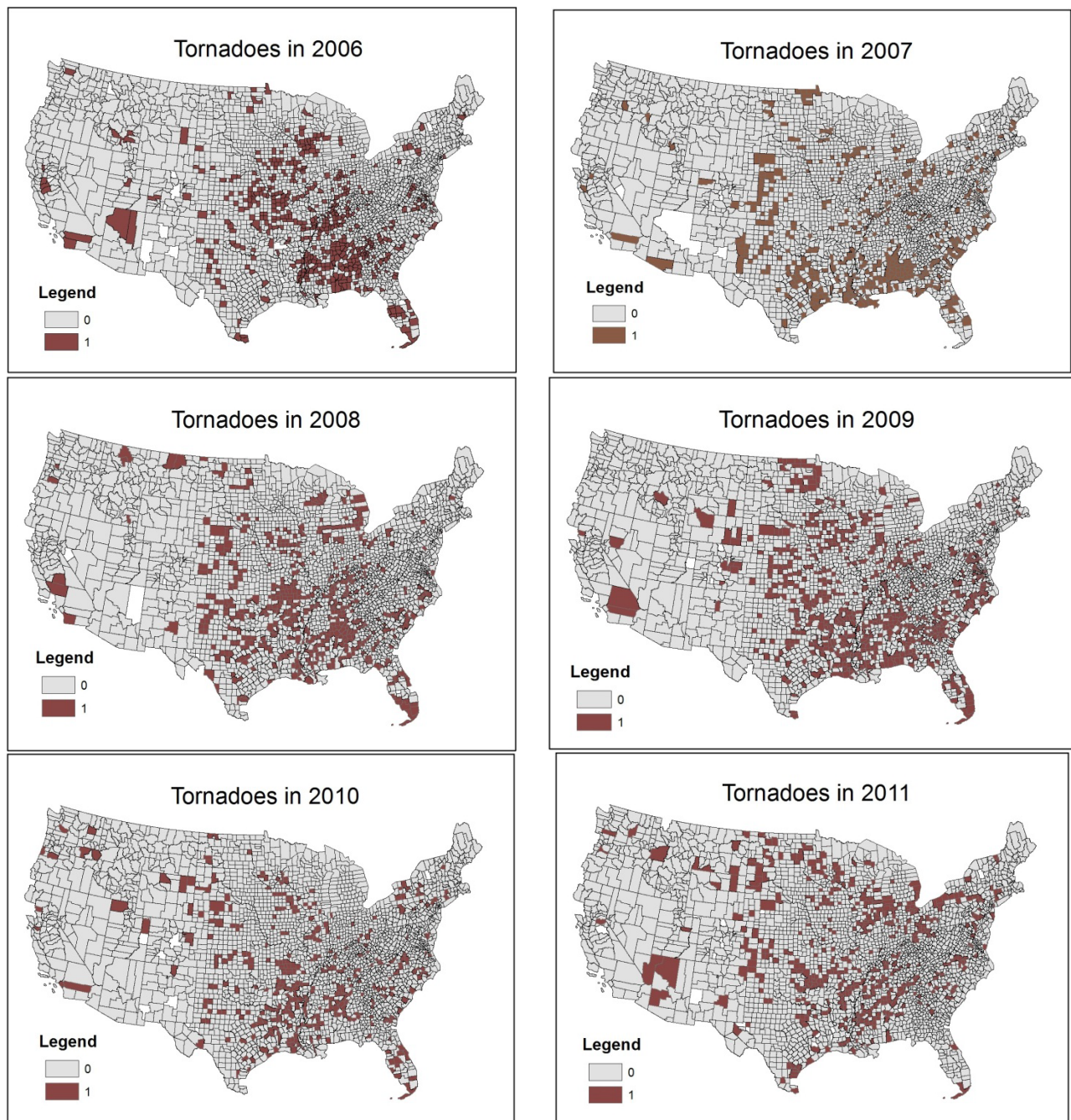


Figure A.1g

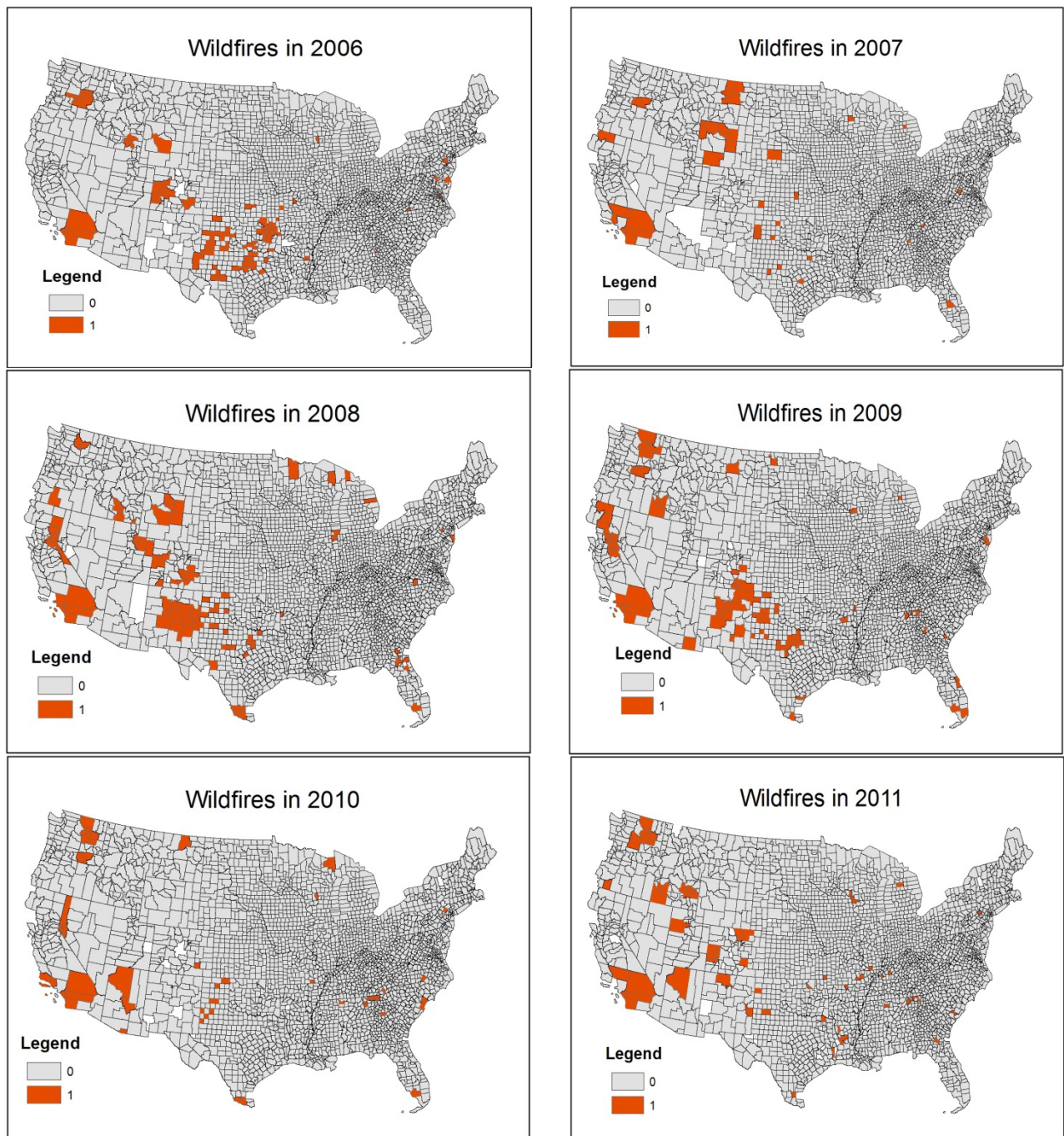


Figure A.1h

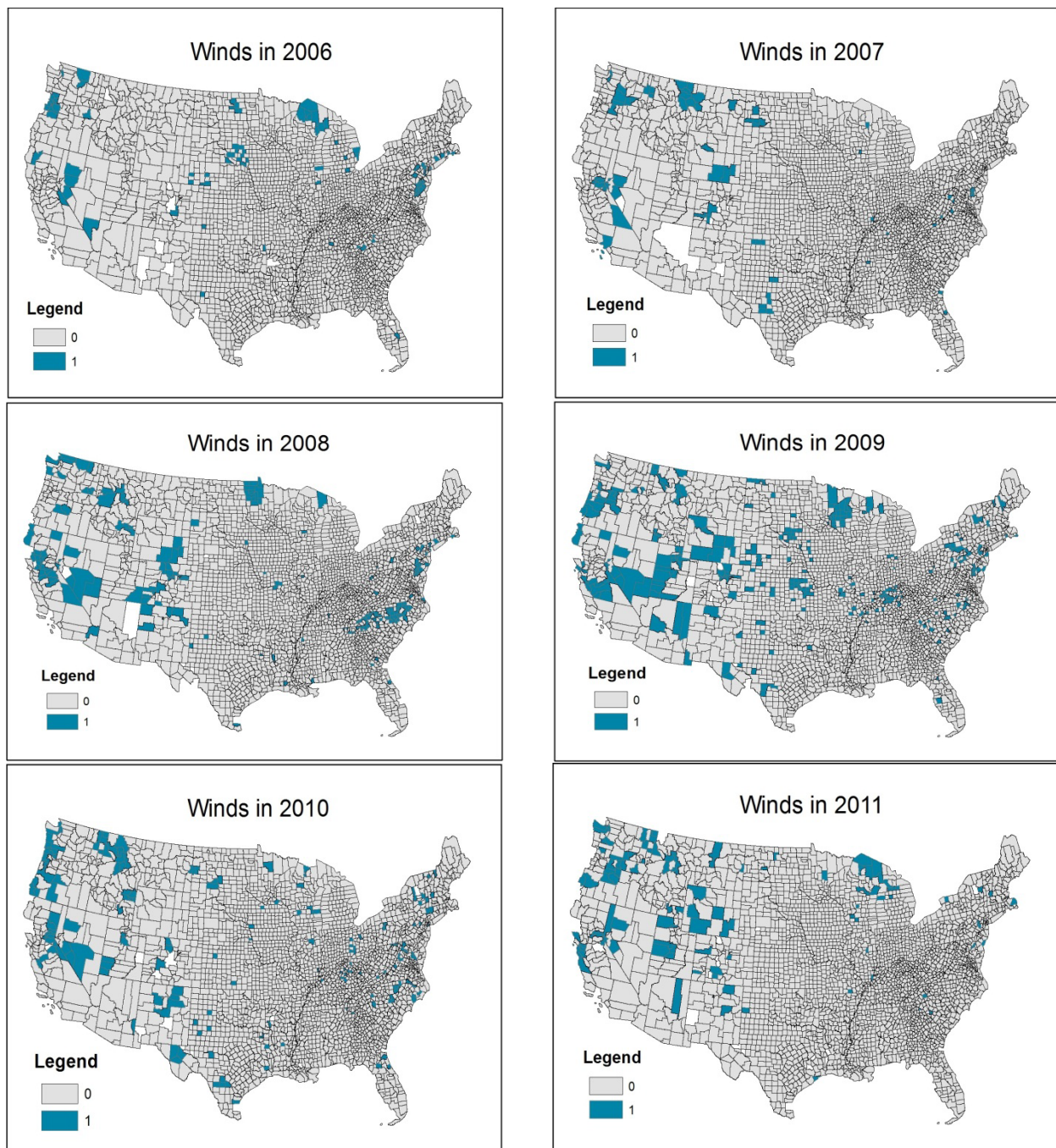


Figure A.1i

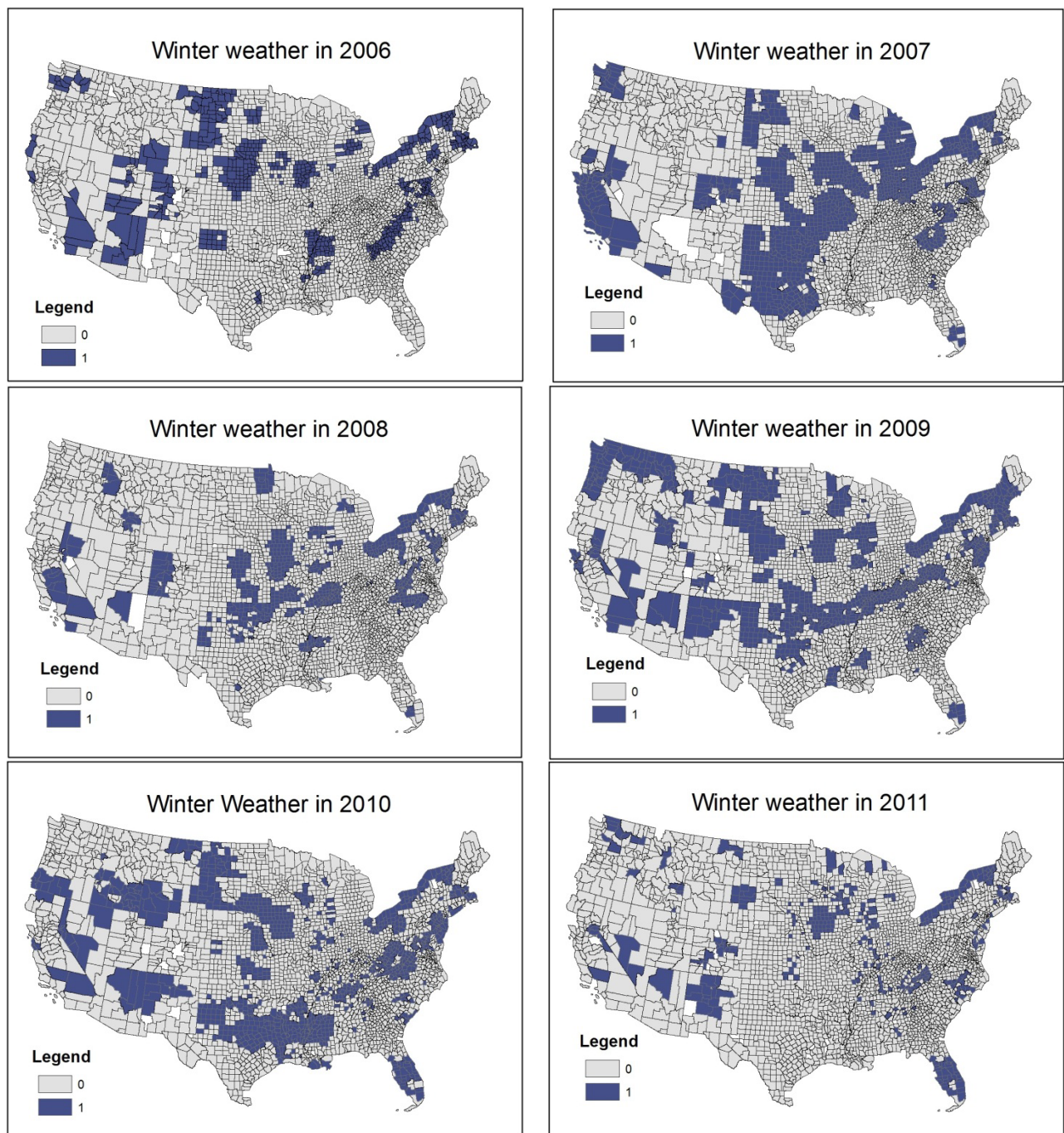


Figure A.1j

Figure A.2

These six maps roughly depict patterns of net out-migration over the years 2005 through 2010, in per-capita terms based on the population of each county in the 2000 Census. Light, medium and dark green counties experienced negative net out-migration (i.e. net in-migration; more people *arriving*). Yellow, orange and red counties experienced positive net out-migration (i.e. more people *leaving*). (Break-points in the positive and the negative halves of the distribution are terciles of the positive and negative distributions aggregated across all years., so colors are consistent for numbers of migrant taxpayers in each year.)

