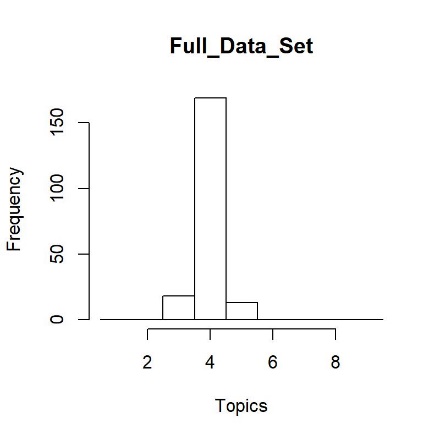
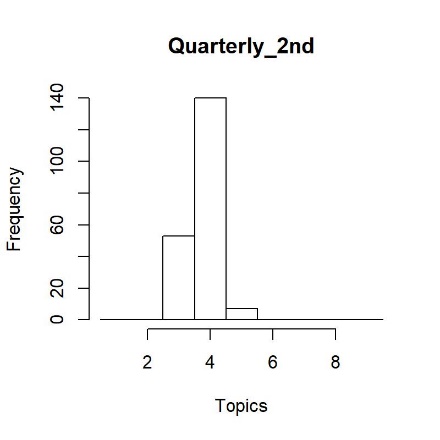
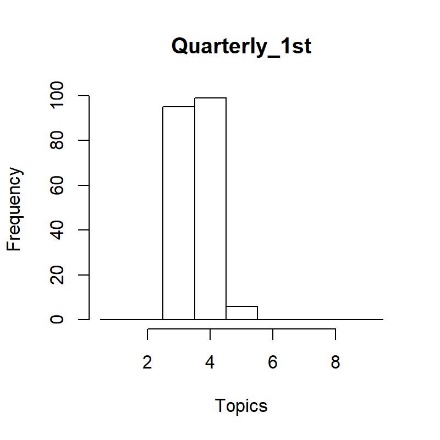
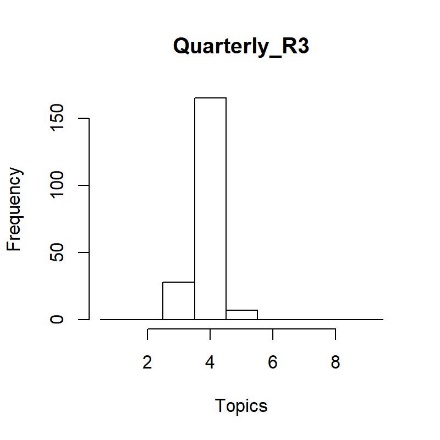
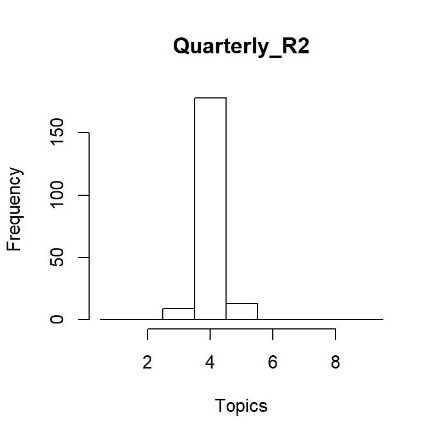
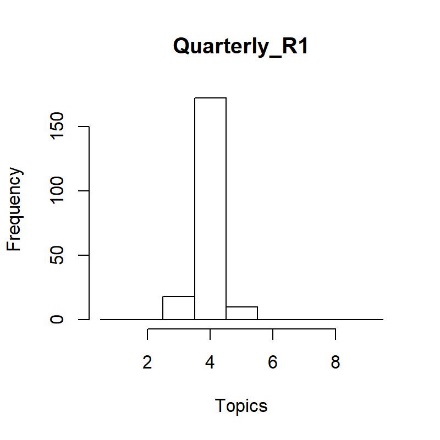
I downsampled the full data set into quarterly (n = 4 per year), semi-yearly (n = 2 per year), and yearly (n = 1 per year) to evaluate the ability of the LDA and change point models to detect patterns with fewer data points. For each of the frequencies, I used both systematic and random samplings. The systematic samplings were the first available and the mid-point (2nd for quarterly, 3rd for semi-yearly, and 6th for yearly) samples. The random samples were random within each time window and I repeated them three times to (somewhat) explore the random variability.

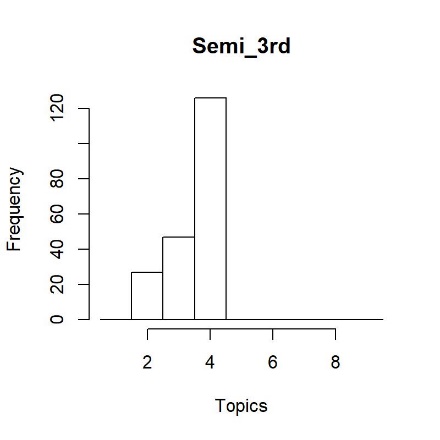
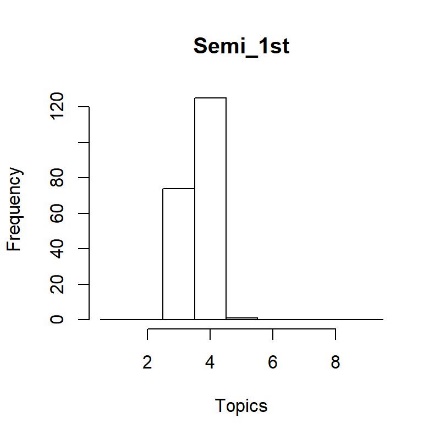
Number of Topics:

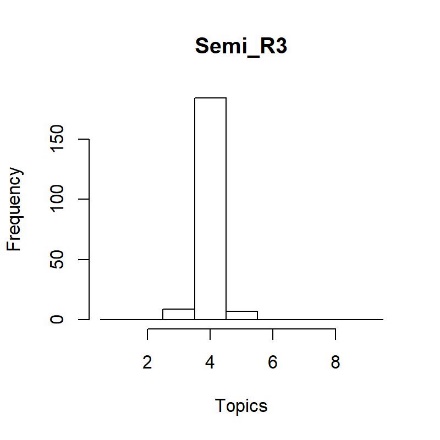
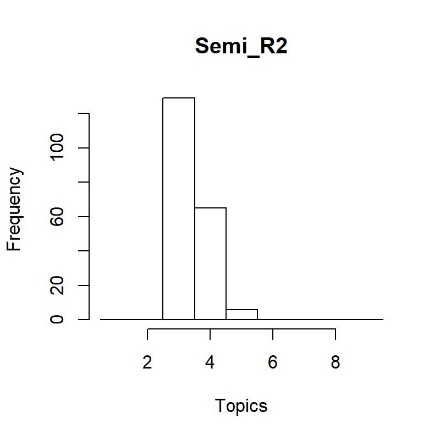
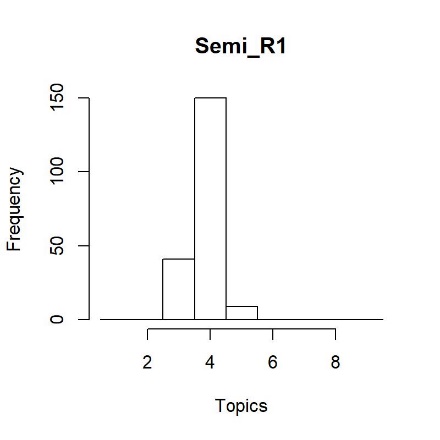
The quarterly data all still pick 4 topics, although the 1st sample within the quarter shows a nearly-similar preference for 3 topics. The semi-yearly data mostly still pick 4 topics, although one of the random samplings picks 3 topics. The yearly data mostly pick 3 topics, although one of the random samplings picks 4 topics.

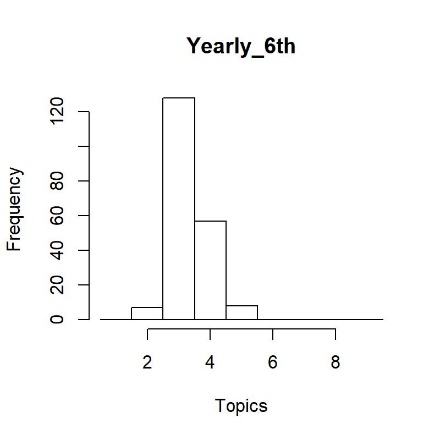
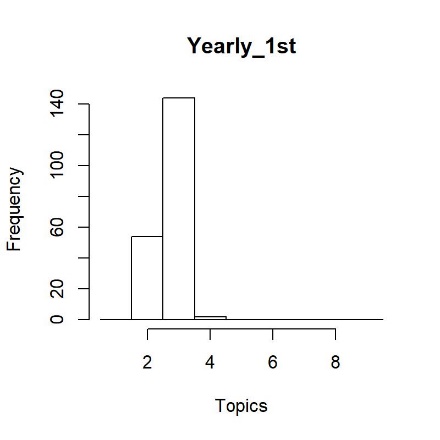


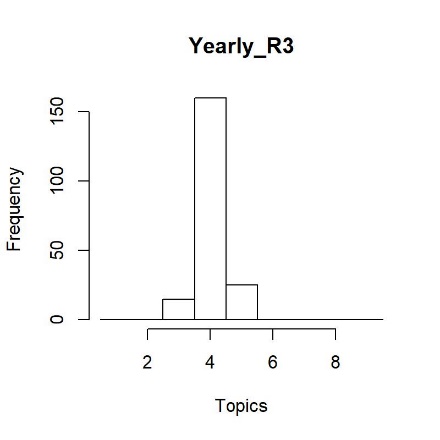
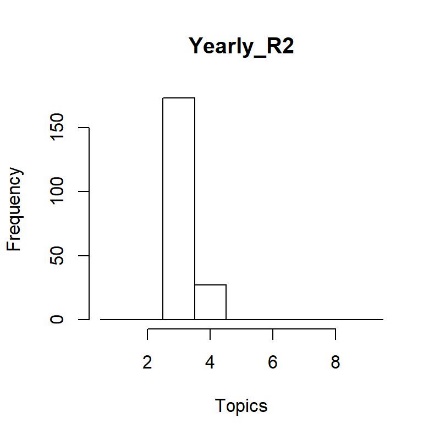
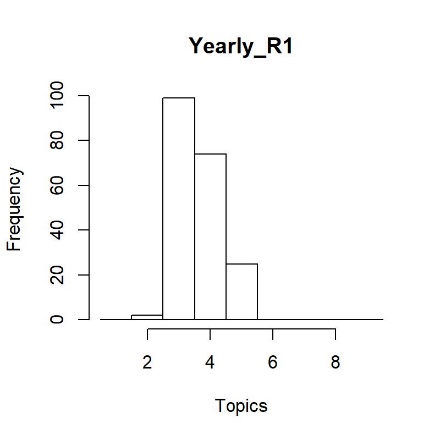












Change points:

As soon as we downsample to the quarterly data, the change point model starts favoring 2 change points. For the quarterly data, these are almost always the 1990 and 2000 change points, but one of the random subsets favored the 1985 and 2000 change points.

The changepoint model favors only 1 change point for the semi-yearly and yearly data. The location of the specific change point chosen moves around a little depending on the specific data used, but it tends to be between 1990 and 2000.

Notes: I ran all of these models using 10,000 iterations in the MCMC. Ideally, we’d run these all with 100,000 or more, but with the NA hang-up we had, I am now running them on my laptop and don’t want to blow it up/wanted to get prelim results together. We currently don’t have a 0 change points model in the suite of models being run, so we don’t have a null model against which to compare the semi-yearly and yearly models (in particular) to see if there is a significant pattern of a changepoint.

