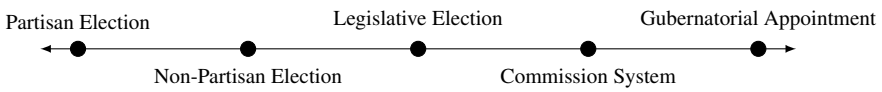
I have specified two different models. A primary specification with five indicators and 42 years of data and a regime change specification consisting of regime change observations over 200 years with four indicators.

Each model was run in R using Martyn **?**'s JAGS and CODA packages for convergence diagnostics **???**.[1] Convergence is assessed using the Gelman-Rubin Diagnostic procedure **?**'s *Superdiag* package in R (**??**).

The primary model is composed of all state year observations between 1800 and 2012. There are four indicators in this model: Appointment Method, Initial Term Length, Subsequent Term Length and Retention Method. There are 8,436 state-year observations.

Figure 1 shows how the results from this model map on to the Independence Continuum discussed in Section **??**

Figure 1: Independence Continuum




Update this with results from model

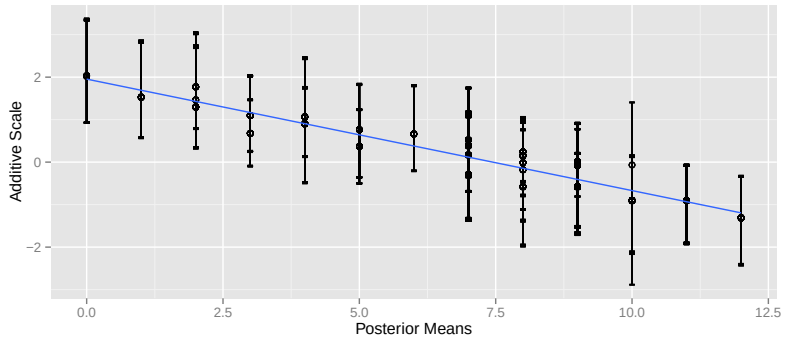
## 0.1 Five Indicator Specification

The Five Indicator specification of this model is composed of all state year observations from 1970-2012. There are five indicators in this specification, the same four in the primary specification, with the addition of the indicator for docket control.

Figure 2 shows the results of an additive scale for the relevant indicators and data in

[1] Other software packages used are listed in the Reference section without being cited in-text.

Figure 2: Five Indicator Specification- Additive vs. IRT




the alternative specification. The Pearson's Correlation between the additive and the IRT model is $-.9553$ with a $p$-value of $< .001$.

Update this with new value

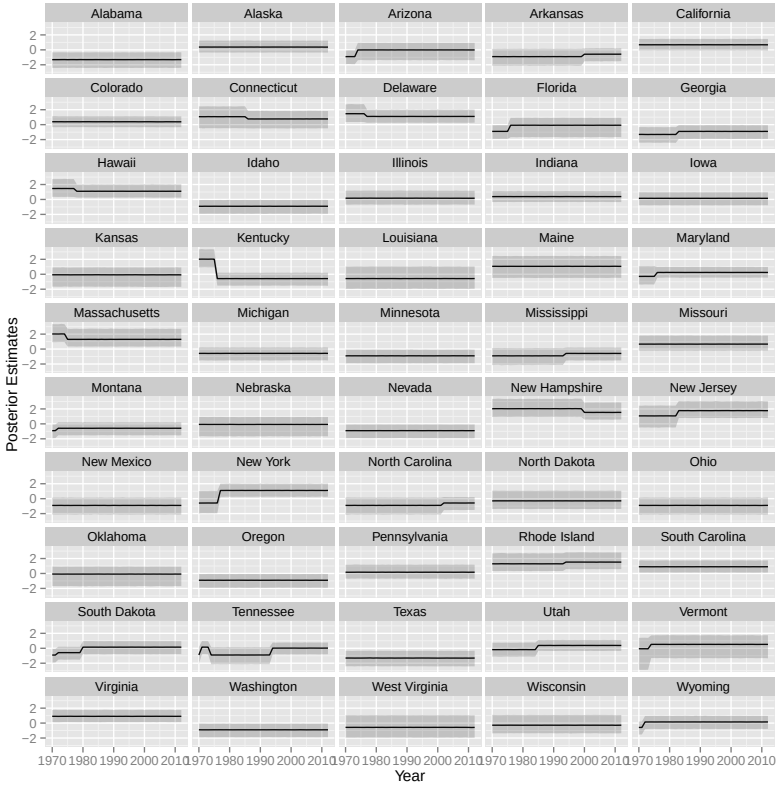
Figure 3 shows the results of the posterior estimates over the time period in the sample. Due to the short time period in the sample, there is little variance in the posterior estimates. However, there is variance between the indicators, which allows for the model to converge. The Regime Change Specification was run using 100,000 draws with the first 25,000 thrown away as burnin. Convergence was assessed with a Gelman-Rubin Diagnostic, with a Multivariate Potential Scale Reduction Factor of (**?**).

Put a Number here for MPSF

## 0.2 Regime Change Specification

The second model specification consists of the four original indicators, but is limited to only unique observations of each judicial institution rather than repeated observations over long periods of time. This specification dramatically reduces the number of observations. For instance, Massachusetts, which has maintained the same selection/retention methods as well as the same term lengths for its entire history has only a single observation. There

Figure 3: *De Jure* Judicial Independence Over time



are 155 state-year observations in this model specification. The Regime Change Specification was run using 10,000 draws with the first 5,000 thrown away as burnin. Convergence was assessed with a Gelman-Rubin Diagnostic, with a Multivariate Potential Scale Reduction Factor of 1.04 (**?**).

Figure 4 shows the results of an additive scale for the relevant indicators in the Regime Change specification plotted against the IRT model. The Pearson's Correlation between the additive model and the IRT model is $-.899$, with a $p$-value of $> .001$.

Figure 4: Regime Specification- Additive vs. IRT

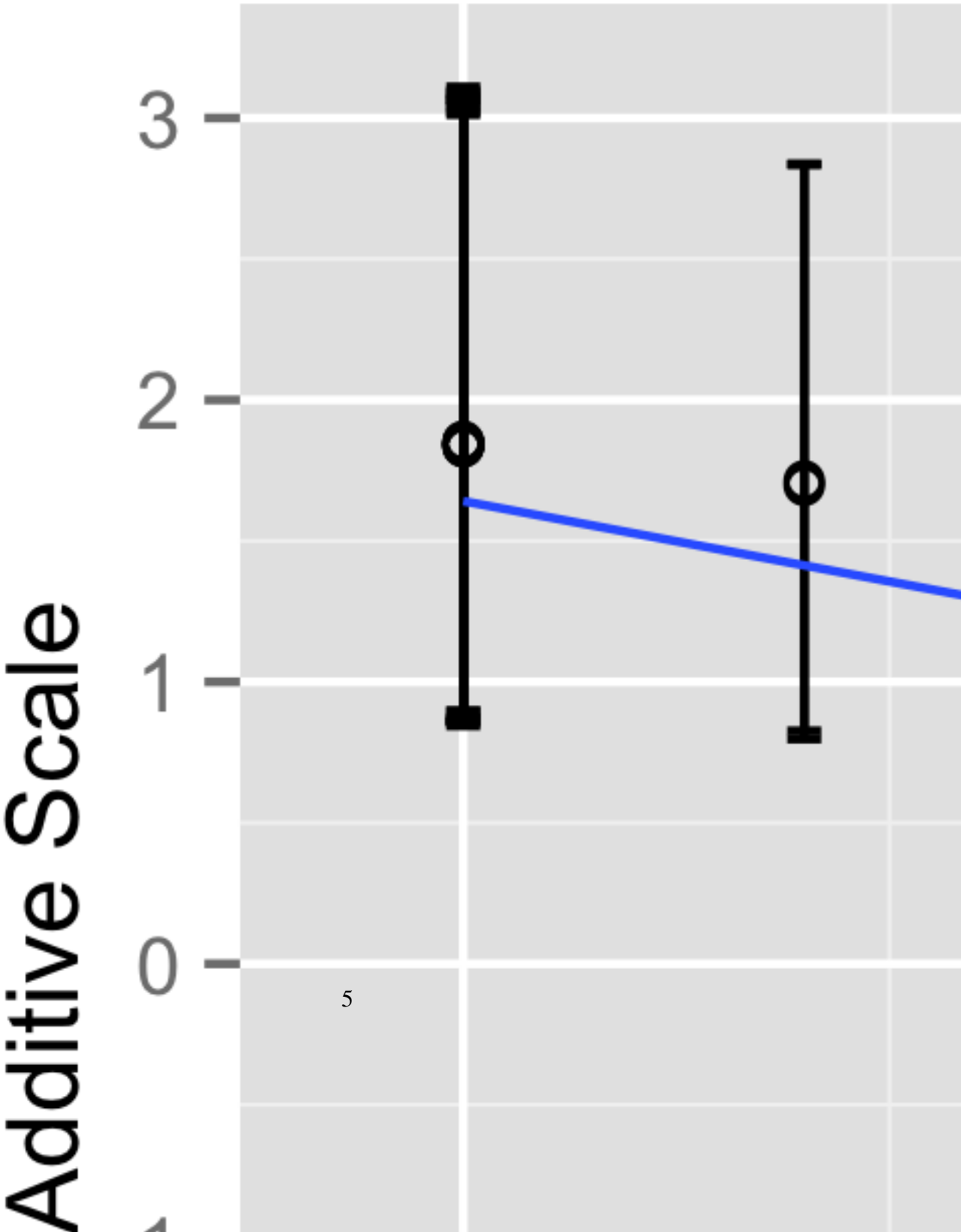


Figure 5: Regime Change Specification- Parameter Means