Modeling Professional Fees with Geographic Effects[143]

To date, empirical research on corporate bankruptcy has largely ignored the reality that any study of bankruptcy involves inherently clustered data: cases are grouped within districts and circuits around the United States. Traditional linear models, such as those presented in the foregoing sections, are based on the assumption that each case is independent, but it is unlikely that cases within districts and circuits are fully independent. Instead, it is likely that they share unmeasured characteristics that make them more similar to each other than to cases in other districts or circuits. Failing to acknowledge and model the hierarchical or nested structure of cases can result in the miss-estimation of standard errors -- leading to an overstatement of statistical significance (*e.g.,* saying that an association is significant when it is not).

Multilevel modeling can be used to account for the hierarchical structure of datasets like the ones used in this study, and to obtain correct estimates of coefficients and standard errors.[144] The exploration of variation between groups (*i.e.* districts), which may be of interest in its own right, is also facilitated by this approach. For example, using a multilevel modeling approach provides an estimate of the amount of variance in fees that is within districts and the amount that is between districts. In addition, it provides an estimate of the range of coefficients across districts. It is possible that associations could be much stronger in some districts or circuits than others. In addition, the multilevel model approach provides a way to appropriately include variables measured at the case or district level in the model at the appropriate level. In the standard multiple regression approach, district level variables (*e.g.,* the indicators of Delaware or New York cases) are analyzed at the case level, but these are in fact district level variables.

One difficulty with using a multilevel model with case data such as this is that the highest level of the model represents the total number of units (or the *n*). That is, running a three level multilevel model on these datasets, with the eleven numbered circuits as the highest level in the

---

143 I am particularly grateful for Julia McQuillan's assistance with this part of the report.

144 For a good, concise discussion of multilevel models, see DOUGLAS A. LUKE, MULTILEVEL MODELING (2004). For a more advanced treatment of MLMs, see STEPHEN RAUDENBUSH & ANTHONY S. BRYK, HIERARCHICAL LINEAR MODELS: APPLICATIONS AND DATA ANALYSIS METHODS (2d ed. 2002).

model, would be similar to conducting regular OLS regression with eleven bankruptcy cases because all of the information from the cases and the districts would be aggregated up to the circuit level. To avoid this problem, I instead opted to construct two level models which examine cases nested within districts. I then address the effects of the 11 circuits by inclusion of dummy variables in some of the models, an approach that is sometimes called “fixed effects” in the econometric literature. That is, I control for circuit effects but do not directly model them.

I examine the final models from both the random sample and the big case dataset using this approach, but have now included the district variables at the district level. One way to conceptualize multilevel models is to imagine running the regression models within each of the thirty-three districts, and then averaging the resulting intercepts and slopes to get the overall estimates. Districts with more cases could be weighted to provide more information for the overall estimates than districts with fewer cases. In addition, we could calculate summary statistics to determine how different district estimates are from each other (called variance components in multilevel modeling). We can then take the thirty-three intercepts and thirty-three slopes, and model the variance in these estimates (deviations from the overall average) using district level predictors. For example, an analysis of the thirty-three intercepts could include indicators of SDNY and Delaware cases to estimate if the average log fees for these districts is higher than in other districts. If individual cases’ characteristics are included in the model, then the intercepts become the statistically adjusted average log requested fees within each district.

The multilevel software creates a separate regression model for each district in the study, and then combines the estimates of the intercepts and slopes using a weighted average to get an overall estimate of the coefficients. Mathematically, the procedure simultaneously integrates the process across all cases and districts, using an iterative maximum likelihood estimation process. The resulting models provide coefficients and standard errors, but because they provide maximum likelihood estimates, the familiar R-square statistics are absent. A baseline model provides estimates of the variance components -- the percent of variance at the case and district levels -- and the full model provides the appropriate coefficients and standard errors. The variance components for each model provide the traditional case level error (level-1, R) and a separate error term for the district level (Uo).

The slopes can also have random effects (individual slopes for each district). Each slope was evaluated separately to see if it did vary significantly across districts in the random sample model. The random sample model includes random effects for professionals3+, committee, log time, log time squared, case dismissed. The big case sample did not include any random slopes. Only the varying intercepts are of substantive interest (i.e, different average log fees between districts), therefore these are interpreted in the results. The chi-square test for the intercept variance components (U0) provides an estimate of the significance of the difference in intercepts (average log fees) between districts, controlling for any case characteristics that are included in the model.

I begin by looking at the basic random dataset model. The results of this analysis are set forth on Table 23. The first model shows that the variance between districts, without controlling for circuits, is significant. About 14% of the variance in total fees is between districts. These differences, however, are explained by the differences in the characteristics of the cases in different districts, as shown by the inclusion of the regression variables from the prior OLS models.[145]

Model 2 shows that compared to all of the other districts, again before controlling for circuits, the SDNY cases cost significantly less and Delaware cases cost more, but this second difference is not significant, controlling for characteristics of the cases. Adding case characteristics significantly improves the fit of the model, is indicated by the large and significant decrease in the deviance statistic.[146] Several case characteristics are associated with total fees. Similar to the earlier models, increases in firm size are associated with increases in cost.[147] Likewise, cases with three or more additional professionals,[148] with committees,[149] higher hourly rates,[150] and first day motions[151] all are associated with higher total costs. Two characteristics -- cases

---

145 The variance component reduces to 0 (100% is explained) and is no longer significant in the next model.

146 Change = 581.7, change in number of parameters = 14, P < .001.

147 B = .111*.

148 B = 1.645***.

149 B = 1.367***.

150 B = .006***.

151 B = .596***.

converted to chapter 7[152] and dismissed cases[153] -- are associated with lower chapter 11 costs.

[152] B = -.992***.
[153] B = -.576***.