Maurer School of Law: Indiana University
Digital Repository @ Maurer Law

IUSTITIA

Volume 4 | Number 2 Article 2

4-15-1977

Using a Prediction Model in Forecasting Appeals

Paul A. Rake

Using a Prediction Model in Forecasting Appeals

PAUL A. RAKE

INTRODUCTION

Decisions involving the expansion of government services are usually of a political nature. However, an accurate understanding of factual data assembled for planning purposes keeps the possibility of measurement available even for the politician. As a consequence, an economic consciousness which has been important to some of the proprietary arms of government has become fashionable throughout. This trend is reflected in the Judiciary by Chief Justice Burger's interest in the collection and measurement of data for analysis in expanding judicial services by creating new judgeships.[1]

As an example, the Federal Judicial Conference makes recommendations for new district judgeships only on a quadrennial basis unless an emergency is indicated by the office's weighted caseload analysis.[2] In a home court event, the Conference found such an emergency in the Northern and Southern Districts of Indiana and recommended legislation to the Congress in 1970. It based its action on statistical data gathered by the Administrative Office of the U.S. Courts.[3] As might be expected, the failure to appoint judges to the Indiana Districts was political — but the accumulation of factual information operated as the catalyst for action and made readily apparent the narrow grounds for the decision.[4]

Many state courts are confronted with situations of emergency proportions which require additional judges or legislatively prescribed changes in procedure. Unlike their federal counterparts, however, state courts rarely have the capacity to continually gather statistical data. Despite political exigencies which constrain most judicial appointments, state courts have little information which can be made available to their legislators as a rational gauge for intelligent use.[5] Thus, the presumption of a political choice is often self-fulfilling because no worthwhile alternative is provided.

Similarly, a lack of adequate data has inhibited state court attempts to plan their daily tasks. In the process, the social function they perform sometimes becomes lost in meaningless detail. Allocation of judge time, distribution of cases and measurement of judicial productivity therefore have a tendency to be

superficial. These factors are especially true for some of the intermediate appellate courts which have recently expanded jurisdictions and consequent rapid growth coupled with inherently limited power to adapt to the change. The need for a planning framework for these courts is critical.[6]

These considerations led the intermediate court of Indiana to attempt to use objective research techniques to provide a plan for growth and adaptation. A planning process and prediction model was developed internally by court personnel for use in a computer program. Altogether, the prediction and planning process has required a minimum amount of time while creating a method of universal potential.[7]

BACKGROUND

Following the 1972 reorganization of the Indiana Court of Appeals into three panels serving defined geographical districts,[8] the Court soon found itself floundering with too many unevenly distributed cases.[9] Lacking a sufficient base of statistical data from which to formulate a plan to cope with the problem, various proposals, including redistricting the court, adding more judges, and developing a more sophisticated staff research, could not be measured for effectiveness or advisability.

In response to these problems, the Court developed a project to deal with the future caseload by constructing a regression model to predict appeals. This model generated estimates of the number of criminal and civil appeals to be filed during 1979, 1980, and 1981 from each county in the state. From these caseload predictions, inferences were drawn regarding the merit of several of the proposed changes. The model has made planning and evaluation for the future a viable function. Moreover, the potential need for more judges can now be statistically justified.

DATA BASE

In focusing on the resources for a prediction model, it should be noted whether one has a data base describing the number of appeals brought before a court during its existence. If so, this population of appeals organized chronologically over a given time period according to type and geography represents the past trend of appellate work. A simple and useful predictive technique is to extend graphically the pattern established by past filings. Variations are then projected into the future on a cyclical basis. Thus, if there were a decline in the number of appeals filed from a particular county, the trend of that decline should continue into the future. Unfortunately, the value of this method varies with the quantity of data assembled in a time series. Since the Indiana Court of Appeals began its operation in 1972, the narrow distribution of data was not susceptible to generalizations which would account for long-term growth patterns.[10]

In order to overcome this inherent weakness in the available data on numbers of appeals, the court developed a different type of model. The hypothesis is that the number of appellate filings is related to various demographic and legal factors in such a way that they can be used as prediction variables in a regression equation. The analysis and control of these indirect variables is easier to use in a forecasting function than is a mere extrapolation of past data based on the number of appeals filed. More importantly, the data available on several types of independent variables over a long period increases the reliability of the forecast.

Thus, with specified degrees of confidence, future workloads can be related to more readily predictable factors with significant accuracy. In this way, the court is able to draw from universal demographic and economic statistics to produce a model for systematically predicting future workloads from various parts of the state. The premise for the Indiana project, therefore, is that appeals can be predicted accurately, a simple model for future use can be developed from available data, and judicial resources can be allocated on the basis of a relevant planning forecast.

VARIABLE SELECTION

Selecting potential independent variables involved determining whether any particular phenomena such as a change in the number of registered motor vehicles varied predictably with a variance in the number of appeals filed. The criteria for accepting or rejecting one of these potential regression variables depended upon its availability, whether it could be easily gathered; its intuitive meaning, whether it would seem to have a causal tendency; its correlation with appeals and its potential for future use. The base year for correlation purposes includes the actual filings of criminal and civil appeals during 1975. Since these variables represent the accumulated data on appellate work for that year, they are used to test the strength of the relationship of the data to the predicting variables chosen. In this context, the effort is to measure how well the value of one of these variables can be estimated on the basis of a knowledge of the other variables.

The initial variables tested included county figures for a variety of demographic variables, such as: population, vehicle registrations, crime levels, traffic flow, land area, judges, law firms, lawyers, employment, income and value added by manufacturing. These figures were gathered from various data sources including the U.S. Census,[11] the Uniform Crime Report,[12] and Martindale-Hubbell.[13] From these original data numbers, certain new variables were generated. For example, from a knowledge of population and land area, population density was developed. Similarly, percentages of lawyers to population, police officers per capita, and a host of other factors were created. Ultimately, over sixty different variables were tested in their capability to forecast criminal and civil appeals.[14]

A comparison of the different variables by county for the 1975 time period to the filings of criminal and civil appeals for that county using the correlation analysis, produced measures of acceptable strength. The statistical measure derives from the computation of the correlation coefficient for the variables involved. Thus r^2 represents the percent of the variability explained in the dependent variable by a change in the independent variable.[15] In accepting the adequacy of the explained variation in a linear correlation, the mathematical test performed showed that the correlation coefficient had a high significance for eleven variables in the criminal appeal category and for ten variables in the civil appeal category.

Correlation Results ($r > .9$) with 1973 data

Criminal

Property damage accidents	Total traffic accidents
Personal injury accidents	Population density
Judges	Urban population
Law Firms	Total number of rated attorneys
A rated attorneys	Funds expended for public defense
B rated attorneys	

Civil

Level of reported crime	Judges
Self initiated police action	Law Firms
Population	Passenger vehicles
Sworn police officers	Total crime level
Vehicle registrations	Rated attorneys

The results of this screening process revealed several important conclusions. First, it identified a multitude of variables which could be useful in predicting appeals. Secondly, it compared the strengths of these variables and disclosed that there were in fact some which could be used with a high degree of confidence. Third, it tested many intuitive hypotheses. Together these attributes confirmed some of the usual presumptions about what contributes to appellate litigation. For example, one might note that the quality and number of attorneys affects civil appeals in the county as do the density of vehicle registrations, and the value added by manufacturing. Similarly, criminal appeals are affected by the density of population, the number of law enforcement officials, and the amounts paid for public defenders. However, not all the usual variables were confirmed for a .90 correlation or better.[16] Thus, while one gets a cursory glance at the causes of litigation, a more important function has been performed by insuring that no presumed factors have been ignored or not properly tested in the context of finding the most reliable predictors of appeals.[17]

Ultimately, the number of highly correlated variables to be used in the regression equation had to be reduced. Only three were introduced into the equation and used in the final prediction model so that the rather high covariances that these independent factors exhibit could not be considered additive in terms of predictive weight. This problem could have occurred because the variables themselves are correlated to each other. Thus, adding more to the function would progressively explain less of the variance. For example, if one independent variable such as the number of attorneys explains 90% of the variance occurring for the change in the number of appeals, an additional independent variable such as the number of judges could not realistically explain more. On the contrary, an equation which includes too many independent variables may be interpreted as exaggerating predictive accuracy since the covariance of these related independent variables does not pertain directly to the variance in the dependent variable. Those variables selected, therefore, were numbers of attorneys, population density and vehicle registration.[18]

EQUATION MODEL

A correlation matrix tests the variability in one set of numbers with the variability in another set. Although mathematically no causal relationship is presumed to exist between the set of variables chosen, it is clear that if the probability of the occurrence of an event is highly correlated with the occurrence of another event, a relationship could exist. In a step-wise regression, the occurrence of events are compared for frequency with the occurrence of other events which have been determined to be related. To formulate this into an equation presumes that the occurrence of appeals is a function of the occurrence of the selected variables. However, since the function explains a special type of relationship, a programmed search is used to find the appropriate coefficients used to minimize the standard error of the estimate. These coefficients produce predictive models for criminal and civil appeals in Indiana with an r^2 of .9734

and .9319 respectively. This means that 97% and 93% of the variations presumed to occur in criminal and civil appeals can be explained by the predicted change in independent variables as disclosed by the functional relationship in the regression equation. Future trends in the predicted data involving population, vehicle registration and numbers of attorneys were then used as the independent variables to determine the number of appeals in the topic years.[19]

RESULTS

Although the total number of forecasted appeals deviated only marginally from what one might have anticipated from past trends, the statistically predicted figures have several additional characteristics. As previously noted, the projected forecasts have an increased reliability based on their relationship to other long run trends. They extend further into the future. They indicate a differentiated pattern of regional growth in the state. Together these factors were particularly useful to the Court of Appeals of Indiana because of its geographic districts. Thus, by comparing average disposition rates for criminal and civil appeals with the number of appeals predicted to arise before any given district, workload factors could be determined for each judge. From these, a plan for realignment of district boundaries to include an additional district with another panel of judges was simulated prior to its recommended implementation.

In a broader context, it should be noted that this type of prediction model can be programmed to suit the particular plans of the individual court. The correlation matrix alone provides a source for finding useful information concerning the relationship between those factors conducive to the litigation of appeals. Presumably, this type of data could be helpful in studying the practices of the bar or in determining the effects of policing the court's rules of procedure. It might also be used to study public defender systems and their effect on appeals. The actual prediction itself would be more useful as a criterion for measuring the effectiveness of a court's planning function. Simulated changes in court operation can be tested. And finally, the administration of caseload assignments to multiple panels can be organized according to future work flows.

The project goal for the intermediate court in Indiana was to provide a plan for growth from an objective research technique. The emphasis on measurable data, empirically derived, complemented the administrative staff function. Political elements which presumably measure different criteria were temporarily put aside so that organizational needs could be attended to. Therefore, a simple model for use in a planning function has, in this case, been effectively used.

FOOTNOTES

[1] Address by Chief Justice Burger, American Bar Association Annual Meeting, Aug. 10, 1970, in H. James, **Crisis in the Courts, (Rev. ed., 1971).**

[2] **Hearings on S. 597 Before the Subcomm. on Improvements in Judicial Machinery of the Senate Comm. on the Judiciary,** 93rd Cong., 1st Sess., at 30 (1973).

[3] **Id.** at 508.

[4] Gary Post Tribune, Oct. 22, 1974, at 20, Col. 1.

[5] Summary, Report of National Center for State Courts on the California Courts of Appeal.
[1975] **Calif. Jud. Council,** at 14.
[1971] **Ohio Jud. Conf. Biennial Rep.,** pt. 4.
[1971-1972] **New York Rep. Admin. Bd. of Jud. Conf.,** at 349.

[6] J. Mueller, **Wisconsin Appellate Practice and Procedure Study** (1975).
D. Halperin, **Report on the Appellate Process in Alabama** (1973).

[7] From a grant of the Law Enforcement Assistance Administration through the Indiana Criminal Justice Planning Agency, computer facilities were made available to the Court by Indiana University under the direction of Mr. Daniel DeHayes, Asst. Prof., Graduate School of Business.

[8] Indiana Constitution, Art. 7, §§5, 6.

[9] IC 1971, 33-2.1-2-2 (Burns Code Ed.)

[10] Two factors militated against a trend analysis: (a) The data base available for a time series projection requires more than a three year interval in order to gain reasonable accuracy for projections in the long run. (b) The smoothing function used by the court required an understanding of the process factors involved in generating appeals not available in a simple extrapolation.

[11] Population estimates for counties were prepared using 1974 official figures published by the U.S. Bureau of Census.

[12] **Uniform Crime Rep. for the U.S.** (1974) (F.B.I. Pub. U.S. Dept. of Justice).

[13] Vol. 4, **Martindale-Hubbell Law Directory** (108th Edition 1976).

[14]

1. 1973 Crim. Appeals
2. 1973 Civil Appeals
3. 1974 Crim. Appeals
4. 1974 Civil Appeals
5. 1975 Crim. Appeals
6. 1975 Civil Appeals
7. 1973 Self-initiated Actions
8. 1973 Prop. Damage Acc.
9. 1973 Per. Injury Acc.
10. 1973 Fatal Acc.
11. 1970 County Pop.
12. 1972 Police Off.
13. 1973 Reported Crime
14. Percentage A Rated Lawyers
15. County Land Area
16. 1972 Non-Urban Traffic Flow
17. 1985 Growth Factor for Traffic
18. 1975 Number of Judges
19. 1973 Number of Law Firms
20. 1973 A Rated Lawyers
21. 1973 B Rated Lawyers
22. 1973 C Rated Lawyers
23. 1973 Lawyers
24. 1973 Passenger Vehicles
25. 1970 Farm Tractors
26. 1970 Negro Pop.
27. 1970 Median Yrs. Ed.
28. 1970 Percentage Unemployed
29. 1970 Median Family Income
30. 1967 Value Added Man
31. 1970 Farm Acreage
32. 1969 Number of Lawyers
33. 1970 Number of Lawyers
34. 1971 Number of Lawyers
35. 1972 Number of Lawyers
36. 1973 Number of Lawyers
37. 1974 Number of Lawyers
38. 1975 Number of Lawyers
39. 1975 Part Time Prosc.
40. 1975 Full Time Prosc.
41. 1975 Defendants Charged
42. 1975 Defendants Tried
43. 1975 Defendants Convicted
44. 1975 Forecasted Pop.
45. 1980 Forecasted Pop.
46. 1985 Forecasted Pop.
47. 1975 S.M.S.A. Status
48. 1970 Vehicle Registrations
49. 1971 Vehicle Registrations
50. 1972 Vehicle Registrations
51. 1973 Vehicle Registrations
52. 1974 Vehicle Registrations

53. 1970 Pop. Density
54. 1975 Judges/Person
55. 1975 Lawyers/Judge
56. 1973 Total Rated Lawyers
57. 1973 Forecasted Pop.
58. 1974 Forecasted Pop.
59. 1973 Pop. Density
60. 1974 Pop. Density
61. 1975 Pop. Density
62. 1975 Funds for Pub. Def.

[15] D. Harnett, **Introduction to Statistical Methods,** 347 (1972).

[16] Generally the failure of correlation for certain variables occurred for two reasons: (a) the intuitive decision to test certain variables was not supported by the empirical data, **e.g.,** per capita income. (b) the number of appeals arising in certain sparsely populated counties was statistically insignificant relative to certain data.

[17] The relative value of predictors depends to some extent on the accuracy of the data made available to the court from a particular county. By comparing the predictability of certain factors, the more uniformly reliable information was ascertained. Moreover, many presumed influences were discounted after their relative values had been determined in the broad context of many different variables.

[18] Within constraints imposed by available data, numbers of attorneys, population density and vehicle registrations were readily available and easily forecasted for long run use. Likewise their correlations were respectively high.

	Attorneys		Population Density		Vehicles	
	Crim.	Civil	Crim.	Civil	Crim.	Civil
1973	.978	.914	.909	.900	.876	.924
1974	.981	.931	.921	.897	.881	.931
1975	.988	.953	.936	.945	.946	.944

[19] A multiple linear regression equation with appropriate partial regression coefficients indicates how much influence each independent variable has on appeal by holding the others constant for each test.