Luskin, Mary Lee "Building a Theory of Case Processing Time"
Judicature 62 (1978) 115-127.

BUILDING A THEORY OF CASE PROCESSING TIME

Mary Lee Luskin

(1978) 62 Judicature 115

The author considers the measurement of "case processing time", rather than "delay", and looks at the possibility of developing mathematical models for the purposes of analysis.

by Mary Lee Luskin

Building a theory of case processing time

Do delay-reduction programs work? We won't really know until we begin constructing analytical models that are appropriate to our theory.

From Shakespeare to Dickens to Warren Burger, delay in court and its evil effects have been decried. Although it has always been a subject of considerable interest, concern with court delay has intensified in recent years. Courts are implementing speedy trial standards and rules and introducing innovations designed to increase court productivity. Attempts are also being made to reduce court caseloads by changing the law and by diverting cases to other forums.[1]

In addition to applied activity, efforts are being made to increase our understanding of

1. Since 1967, for example, three sets of standards and goals for processing criminal cases have been proposed. *See* American Bar Association Project on Minimum Standards for Criminal Justice, STANDARDS RELATING TO SPEEDY TRIAL. Chicago, 1968; National Advisory Commission on Criminal Justice Standards and Goals, COURTS. Washington, D.C., 1973; and The President's Commission on Law Enforcement and the Administration of Justice, TASK FORCE REPORT: THE COURTS. Washington, D.C., 1967.

In addition, The Speedy Trial Act of 1974 (see U.S.C. §§161–74 [1974]) and speedy trial provisions in 44 states set time limits for events in the criminal process. No fault insurance laws, proposals to change the jurisdiction of the federal courts, neighborhood justice centers, and pretrial diversion programs have all been justified—at least in part—on their promise to reduce caseloads.

Even more numerous are proposals for structural, procedural and managerial innovations intended to reduce court delay—e.g., hire a court administrator, require omnibus hearings, write fewer opinions, unify state trial courts, adopt an individual (or master) calendar, and abolish the grand jury.

the causes of court delay. The National Institute of Law Enforcement and Criminal Justice—the major governmental source of funds for criminal justice research—has made court delay one of its priority research areas. Current and just completed research on court delay includes a study of the impact of state speedy trial provisions and studies of state courts and federal district courts to develop measures of delay and to assess the impact of procedures on delay.[2]

Despite this activity, we have relatively little understanding of how court delay is related to demand for court services, to available court resources, to procedures used to settle disputes, and to the motivations of participants. To gain this understanding, we need to develop and test a theory of case processing time.

Case processing time

I choose the words "case processing time" rather than "court delay" with care. Delay usually denotes abnormal or unacceptable time lapse. Most would agree, however, that some time used in processing cases is normal, acceptable and even desirable. Yet the measures of court delay that researchers invariably use (e.g., median days from arraignment to final disposition in criminal cases or median days from service of the answer to trial in civil cases) are really measures of case processing time, that is, normal *plus* abnormal time lapse, and not of delay alone.

To measure case processing time and call it delay has disadvantages. First, it is inaccurate. Second, in calling all case processing time "delay," we tend to load value choices that may be necessary between speed in case processing and other goals in favor of speed. Third, the use of the term delay, when case processing time is meant, may suggest that the causes of abnormal case processing time are distinct from those that produce normal case processing time. It may restrict the range of variables we consider as causes, thereby limiting our explanation. The use of the term "case processing time" rather than the pejorative "delay" for the period in question is, I think, both more accurate and theoretically productive and will help us recognize competing goals.

What we need, then, is a theory of case processing time that explains variation across individual court cases, jurisdictions and time. The duration of a criminal case, for example, can be less than a day if the defendant pleads guilty and is sentenced at arraignment. Other cases can take weeks or months to complete. We expect that variation in the amount of time taken to process individual cases is not random but, rather, that there are systematic differences among those cases that take more or less time.

Case processing time also varies across jurisdictions. For example, the median number of days between arrest and trial courtroom disposition for felonies in 1972 was approximately 268 days in Chicago and 226 days in Baltimore, but only 71 days in Detroit.[3] We expect community, court, and personnel characteristics to be associated with differences like those in these three cities. Finally, case duration varies over the history of a court or a society. Once again, we assume the variation is systematic. We expect it to be associated with changes in the law and the courts and with the reactions of court personnel to these changes. A comprehensive theory of case processing time should specify and estimate the effects of all of these sources of variation.

In this article, I want to consider the contribution of past literature and research on court delay to the development of a theory of case processing time. My purpose is to indicate gaps in our knowledge rather than to summarize what we know. I shall focus, therefore, on the methodological limitations of previous work and the resulting defects in our understanding of case processing time rather than the considerable contributions and strengths of this work. By doing so, I hope to demonstrate the need for a comprehensive theory of case processing time and to argue for the research designs and statistical analysis models that I believe

2. These projects are being conducted by the Midwest Research Institute, the National Center for State Courts, and the Federal Judicial Center, respectively.

3. James Eisenstein and Herbert Jacob, FELONY JUSTICE: AN ORGANIZATIONAL ANALYSIS OF CRIMINAL COURTS 233. Boston: Little Brown and Co., 1977.

are most conducive to the development of this theory.

I. Court delay literature

Writing on court delay is voluminous, but much of it might simply be termed inspirational. Judges, administrators and prosecutors who feel that they have been successful in reducing delay in their own courts describe their actions and report "significant" reductions in delay or backlog.[4] The authors assume the actions have caused the reduction in delay but make no attempt to actually separate the effect of the innovation from that of other potential causes nor to measure the size of the effect that can be attributed to the action taken by the court relative to other variables. In the absence of a theory of case processing time, these reports may be helpful to court administrators, judges and prosecutors looking for suggestions about how they might approach problems of court delay, but they contribute little or nothing to the development of a theory of case processing time.

More analytic works on court delay specify variables thought to affect case processing time and backlog and describe the relationships among them. The articles discussed below vary in the degree to which they do this, in appropriateness of the statistical models used to represent their theory, and, thus, in their contribution to a theory of case processing time. I have organized this work under four general headings: (1) inventories of causes, (2) experiments, (3) surveys, and (4) process models.

A. Inventories of causes

Judges, lawyers and other close observers of the courts have provided us with a rich set of hypothesized causes of and remedies for court delay.[5] The difficulty with inventories of causes as theory, however, is that they do not describe the relationships among the causes. We cannot tell from a list whether an independent variable (a cause) affects the dependent variable (delay or case processing time) directly or through its influence on one or more intervening variables.

For example, one inventory of causes of court delay lists, among other variables, "the darkened courtroom," and the propensity of certain attorneys to take more cases than they can handle as causes of court delay.[6] One might hypothesize that concentration of the trial bar leads to a greater probability of scheduling conflicts, which, in turn, lead to underuse of courtrooms (presumably an indicator of low judicial productivity); but until such an ordering is made for these and other variables, we have a theory only at the most elementary level.

In addition, inventories of causes of court delay have been limited by lack of agreement on an operational definition of delay. Such diverse phenomena as number of cases pending, ratio of current cases to terminations, time lapse between processing stages (e.g., arrest and trial) in individual cases, median case processing time, and percentage of "overage" cases on the docket have been used as indicators of delay.

Most inventories of causes of delay fail to specify which indicator is being used, even though different indicators (e.g., ratio of current cases to terminations versus time lapse between arrest and sentence) imply different levels of analysis (court-wide versus individual) and may imply different sets of causes. In some lists, the dependent variable (e.g., median case processing time or case processing time in individual cases) can be inferred from the independent variables (e.g., strong versus weak continuance policy

4. *See, e.g.* Ruggero Aldisert, *A Metropolitan Court Conquers its Backlog. Part II: From Pure Pre-trial to Compulsory Settlement Conferences*, 51 JUDICATURE 247–52 (1968); Edward J. Blake and Larry Polansky, *Computer Streamlines Caseload at Philadelphia Common Pleas Court*, 53 JUDICATURE 205–9 (1969); Robert F. Leonard, *Deferred Prosecution Program*, 8 THE PROSECUTOR 315 (1973); and Edward Thompson, *Backlog Disappears or Counselor 'You are Ready'?* 13 COURT 2 (1974).

5. Examples are Richard S. Miller, *Program for the Elimination of the Hardships of Litigation Delay*, 27 OHIO STATE L.J. 402 (1966); State of Massachusetts, Governor's Select Committee on Judicial Needs, "Report on the State of the Massachusetts Courts," (Boston, 1976); Maurice Rosenberg, *Court Congestion: Status, Causes, and Proposed Remedies*, in Harry W. Jones, ed., THE COURTS, THE PUBLIC, AND THE LAW EXPLOSION. Englewood Cliffs, N.J.: Prentice-Hall, 1965; and G. S. Reynolds, *Court Delay Delineated*, 12 WASHBURN L.J. 12–27.

6. Miller, *supra* n. 5, at 407, 410.

or bail status of the defendant) but, in other lists, the independent variables imply more than one dependent variable and more than one level of analysis. Even when this is not the case, the failure to be conceptually clear undermines the theoretical utility of the lists.

Inventories of causes provide a checklist of variables that might be included in a theory, but they themselves do not constitute a satisfactory theory of case prosessing time. In the early stages of theory building, it may be difficult to do much more than list possible causes, leaving the tasks of sorting out levels of analysis, of specifying the causal ordering, and of specifying the form of the relationships among the variables to those who follow. The listing of causes should now be essentially complete; it is time to address the subsequent tasks.

B. Experiments

A few studies of court delay have utilized experimental designs. Experiments are powerful research tools. Their main advantages are certainty of the causal ordering (the independent variables are manipulated and then the dependent variable is observed) and, through physical controls and randomization, high assurance that the hypothesized causes are not confounded with any unmeasured causes.[7]

In an experiment to study the effect of pretrial conferences on case processing time, for example, cases might be randomly assigned to two groups. The cases in one group would receive pretrial conferences, while those in the other would not.The random assignment would make it highly likely that the groups of cases assigned to each treatment were alike on other variables—like case complexity—that might affect case processing time. The inference that any observed effect is due to the experimental variable (e.g., pretrial) and not to other, unmeasured variables (e.g., case complexity) would be strong.

7. For more detailed discussions of the strengths and weaknesses of various research designs, *see* Fred N. Kerlinger. FOUNDATIONS OF BEHAVIORAL RESEARCH: EDUCATIONAL AND PSYCHOLOGICAL INQUIRY 359-408. New York: Holt, Rinehart, and Winston, 1964.

Some courts may not think it is fair to experiment— to try new procedures for some cases but not others.

However, there are ethical and practical disadvantages to field experiments. Experimenters must be able both to manipulate the independent variable and to have control (either for purposes of matching or randomization) over which cases receive which treatments. Court participants may be unwilling, out of a sense of fairness or because of extra cost or work involved, to allow new procedures to be used in some cases and not others. For example, the Vera Institute's Manhattan Court Employment Project offered participation to all eligible defendants arraigned on weekdays rather than randomly assigning them to participation/nonparticipation groups because the project staff felt that ". . . the relationship between a committed staff member and participant would be too close to permit us to withhold services from selected participants for research purposes."[8]

Maurice Rosenberg encountered a similar problem (but for different reasons) when he attempted to measure the impact of pretrial conferences. Instead of random assignment to the pretrial/no pretrial conditions, the court decided to hold a pretrial when either

8. The Vera Foundation. *The Manhattan Court Employment Project/ Summary Report on Phase One: Nov. 1, 1967 to Oct. 3, 1969* at 66. New York: Vera Institute of Justice, 1970.

9. Maurice Rosenberg, THE PRETRIAL AND EFFECTIVE JUSTICE: A CONTROLLED TEST IN PERSONAL INJURY LITIGATION 19. New York: Columbia University Press, 1964.

C. Surveys

Surveys are by far the most common research designs for studying court delay. In surveys, unlike experiments, the researcher does not manipulate the independent variables; both the independent and dependent variables vary naturally. The researcher observes variation in the dependent variable, hypothesizes causes for it, and seeks to sort out and estimate the effects produced by these causes.

This characteristic lack of control over the independent variables means that the causal ordering of the variables is more problematic in surveys than in experiments. Although it is clear that the original charge filed against a defendant is causally prior to the number of motions filed in the case, it may be unclear whether change in caseload causes change in case processing time or change in case processing time causes change in caseload. Lack of control over the independent variables also means that statistical techniques must be used to disentangle the effects of multiple causes of the dependent variable.

But survey designs possesses several advantages. First, they are often feasible where, for ethical or logistical reasons, experiments are not. They also tend to have greater naturalness since the researcher intrudes less into the process he studies. And, third, because it is difficult to manipulate a large number of variables in an experimental design, researchers using survey designs are more likely to be able to obtain estimates of the effects of multiple causes of the dependent variable.

Often, however, the theoretical and policy utility of court delay studies using survey designs has been limited by the inadequacy of the statistical models used to test the researcher's hypotheses. Most researchers hold, at least implicitly, theories of case processing time that are multivariate and asymmetric, that is, theories that say that variation in case processing time is produced by at least two independent variables and that influence flows *to* the dependent variable *from* these variables (e.g., defendant's bail status and the type of charge against him affect case processing time). The analysis models many of these same researchers use, however, are bivariate (they examine the association between two variables only) and symmetric (the statistic estimates the degree to which the independent and dependent variable vary together, but not the change in the dependent variable produced by the change in the independent variable). If the theory is multivariate and asymmetric but the analysis model is bivariate and associational, the model is misspecified.

It is difficult to analyze the effects of *court-wide* variables on case-processing time because of the high cost of gathering data.

Specification error can result from the omission of a relevant variable, the inclusion of an irrelevant variable, and the incorrect mathematical representation of the relationships among the variables included.[11] (A bivariate, associational analysis model for a multivariate, causal theory contains two of these sources of specification error; relevant variables are omitted, and the relationships among the included variables are incorrectly represented.) The consequences of specification error vary with its sources, but failing to include a relevant variable (a common specification error) results in biased estimates of the effects of those variables that *are* included. From a seriously mis-specified model (e.g., one that omits an important causal variable), one can develop neither satisfactory theory nor satisfactory policy recommendations.

11. For a fuller discussion of specification error, see Jan Kmenta, ELEMENTS OF ECONOMETRICS. New York: Macmillan, 1971. Kmenta considers a number of sources of specification error and examines their consequences at 391–405.

Some examples will illustrate the shortcomings of a bivariate, symmetric model for analyzing case processing time. In their study of continuances in the Cook County Criminal Courts, Banfield and Anderson found associations between case characteristics—e.g., prior record, attorney type, and bail status—and number of appearances for a sample of defendants.[12] They arrive at conclusions about these associations by comparing cross-tabulations of two variables with a control for a third. (In one or two of their tables, they add a second control.) This cumbersome approach makes it difficult to take account of more variables simultaneously, but without doing so, they cannot tell us the size of the impact of any one variable given the others.

Similarly, Sovern and Rosenberg looked at the relationships between case characteristics and case processing time in civil cases.[13] They reported that the size of the claim has the single largest correlation with case processing time. Because they measure bivariate associations without having introduced controls for variables that affected the size of those associations, they did not rule out the alternative hypothesis that the association is spurious—that is, that some other variable produced the observed association between claim size and case processing time. Furthermore, if the relationship is real, do we want to know only that size of claim and case processing time vary together? Or do we want an estimate of the magnitude of the effect of size of claim given the presence of the other causes? If we believe that differences in the size of the claim cause differences in case processing time, a symmetric statistical model is inappropriate to our theory and does not yield the information we want.

Effects of court-wide variables

Because of the cost of gathering data on a sufficient number of courts to make rigorous analysis possible, analysis of the effects of court-wide variables has been particularly difficult. James France, for example, compared case processing times in seven jurisdictions that varied on such court level variables as number of judges per capita, individual versus master calendar, and presence of a court administrator.[14] Because the comparisons he makes are mostly bivariate—that is, he looks at the relationship between a single independent variable and the dependent variable—he cannot assess the relative effects of variation on each of the dimensions he measured nor the effects of other dimensions on which the jurisdictions varied.

The lack of control means that causation is unclear. France found, for example, that courts without professional administrators tended to process cases faster than those with administrators.[15] Is he to conclude that administrators slow courts down? Certainly alternative hypotheses exist. Another unmeasured cause may be producing both high case processing time and the tendency to hire a court administrator. Perhaps processing time in the courts with administrators would be even higher without them.

France's analysis, which includes only a few of the potentially important variables—the presence of the administrator and a rough measure of community size—is insufficient to rule out plausible alternative hypotheses. He can note an association between the presence of a court administrator and higher processing times, but the mere existence of such a bivariate association for a few jurisdictions does not provide much theoretical or policy guidance. Alternative hypotheses have not been ruled out and, even if one were convinced that having a court administrator did cause longer case processing times, the method does not provide an estimate of how *much* longer.

Martin Levin compared median case processing times in five criminal courts that vary in caseload and in court practices. Comparisons across only five courts means

12. Laura Banfield and C. David Anderson, *Continuances in the Cook County Criminal Courts*, 35 U. CHI. L. REV. 276 (1965).

13. Michael I. Sovern and Maurice Rosenberg, *Delay and the Dynamics of Personal Injury Litigation*, 59 COLUM. L. REV. 1115 (1959).

14. James France, *Judicial Administration: The Administration: The Williamsburg Consensus—Some Errors and Omissions*, 14 WILLIAM AND MARY L. REV. 1–45 (Fall, 1972).

15. *Id.* at 27.

that hypotheses must be evaluated informally. Levin writes that "if all other factors were equal we might find the larger the caseload, the longer the delay."[16] The "other factors" that are not equal are continuances, motions and full length trials. These court practice variables are also hypothesized to affect case processing time, but Levin's method—dictated by the small number of courts—prevents him from estimating the size of the effect of caseload given these variables.

Although Levin cannot estimate these effects, his work is important because of the theoretical connection he makes between court practices associated with longer case processing times and the motives of participants. Often, the impression conveyed by discussions of case processing time is that caseflow management is isolated from other issues in the criminal court process. Levin's work emphasizes the attitudes and motivations of the participants. A comprehensive theory of case processing time must take these variables into account.

Measuring all the variables

The studies described thus far have measured only subsets of variables thought to affect case processing time. In order to obtain accurate estimates of the effect of any one variable, all of the relevant variables need to be measured and included in the estimating procedure. A National Center for State Courts Project collected data on a wide range of variables that can affect case processing time in state trial courts.[17] Twenty-one courts were surveyed initially and eight were chosen for more intensive study. In each of the eight, cases terminated in 1976 were sampled. Data collection has been completed on this project, but the analysis has not yet been reported. The size of the samples from each court (500 criminal and/or 500 civil cases) makes possible multivariate analysis of the effects of case characteristics on processing times within each jurisdiction.

The District Court Studies Project also seeks to relate differences in speed and productivity to differences in the practices of courts.[18] The project has two goals: to identify differences between "slow" courts and "fast" courts and to identify the ways in which current measures of speed and delay may be misleading. To accomplish these goals, five metropolitan federal courts representing the extremes in speed and delay in the federal system were studied. Once again, the design looks at a small group of fast and slow courts that vary on many dimensions. Because of the small number of courts studied and the cross-sectional nature of the data, most of the analysis relies on informal observations of associations between various activities—opinion writing, for example—and measures of speed or productivity.

There are a few exceptions to this analysis strategy in the District Court Studies Project. One of these is the comparison of two of the district courts in the context of an analysis of variance model in which differences in processing time between two courts are attributed to differences in practices.[19] This analysis permits the addition of some control variables, but the statistical model does not allow the researchers to ascertain the size of the effect attributed to a particular variable on processing time.

In the reports on this project, the informal analysis model used, for the most part, is bivariate even though the implicit theoretical model is multivariate. As a result, the authors cannot assess the relative importance of differences among courts in the speed of processing nor can they rule out the possibility that the relationships they observe are spurious. The District Court Stud-

16. Martin Levin, *Delay in Five Criminal Courts*, 4 J. LEGAL STUDIES 83 (1975).

17. National Center for State Courts, Pretrial Delay Project: Research Design. San Francisco: NCSC, National Conference on Metropolitan Courts, 1977. The description of data collection and analysis that follows in the text is based on personal communication with Thomas Church, principal investigator of the project.

18. Steven Flanders, CASE MANAGEMENT AND COURT MANAGEMENT IN UNITED STATES DISTRICT COURTS. Washington, D.C.: Federal Judicial Center 1977.

19. Steven Flanders and Alan Sager, *Case Management Methods and Delay in Federal District Courts*, in Russell Wheeler and Howard Whitcomb, ed., JUDICIAL ADMINISTRATION: TEXT AND READINGS. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1977.

ies Project can point to some variables that may be related to delay and may be susceptible to manipulation, but it cannot assess the size of the effects to be expected.

Economic models

The analysis models described thus far are primarily associational. The best of this research does separate the effects of some independent variables, but since most of them do not include all of the variables the researcher thinks are important, the models are mis-specified, although the informal nature of much of the analysis tends to disguise this fact.

In contrast, studies of court delay by economists have used multivariate techniques for multivariate theory.[20] These researchers have employed concepts and relationships derived from economic theory to explain the behavior of court actors and outcomes from courts. Most economic models of the courts have not been concerned with explaining case processing time, but rather include case processing time as an independent variable in explaining other variables. In these models, case processing time is conceptualized as a non-monetary price that rations demand for court services.

For example, William Landes hypothesized that demand for criminal trials is a linear, additive function of the per cent of defendants on bail, the average wait for a trial, the average wait for a settlement, the average sentence after conviction by trial, and the average cost difference between trial and settlement. The equation representing this hypothesis is estimated for federal district courts and for a sample of state courts.[21] (The estimating technique, ordinary least squares, is appropriate to the model.) Estimates of the *size* of the effects of each of the independent variables given the others are obtained. This information is much superior to that obtained from associational models.

Data limitations, however, prevented Landes from estimating the full equation for either data set. Neither the federal district court data set nor the state court data set contained information on all of the variables in the equation. As a result, each estimated model is mis-specified and the estimates of the model's structural coefficients are biased. Although the best means of overcoming this problem would be to collect data on all of the variables in the model rather than estimating the model from data sets collected for other purposes, this solution would require a large investment of research resources. Landes's solution—to estimate the models for two existing data sets that differ in the variables missing—does permit some idea of the amount of bias introduced by the omission of variables from the equations estimated. It also provides first approximations of the size of the effects of the independent variables on the dependent variable.

Although flawed, the use of a multivariate, causal data analysis model to represent the theory is superior to the alternative of performing a less rigorous analysis. Using measures of association, for example, would, in addition to the error caused by the missing variables, misrepresent the hypothesized form of the relationship (i.e., associational rather than causal). It is also likely that further specification error would be introduced by the inability to deal with all of the independent variables simultaneously. The presence of this error, however, would be less obvious because the less rigorous techniques do not force the researcher to be explicit about which variables are included in the theoretical model or about the functional form of the relationships among them.

The analysis model Landes uses keeps the shortcomings of the actual estimation in clear view. Doing so not only helps in evaluating the research in question, but also points out what needs to be done to improve upon the theory, its representation, or the estimation of the parameters of its represen-

20. *See* Robert W. Gillespie, JUDICIAL PRODUCTIVITY AND COURT DELAY: AN EXPLORATORY ANALYSIS OF THE FEDERAL DISTRICT COURTS. Washington, D.C.: Government Printing Office, 1977; William Rhodes, *The Economics of Criminal Courts: A Theoretical and Empirical Investigation*, V. J. LEGAL STUDIES 311 (June, 1976); and William M. Landes, *An Economic Analysis of the Courts*, 14 J. LAW AND ECONOMICS 61 (1971).

Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, II J. LEGAL STUDIES 399 (1973) is a purely theoretical treatment of delay and other topics from an economic point of view.

21. Landes, *supra* n. 20, at 78.

tation. It becomes more likely that theory and research will be cumulative.

An illustration of a researcher building upon earlier work is William Rhodes's extension of Landes's model. Rhodes extends that model to take account of the effect of case processing time on demand for trials. He hypothesizes that case processing time affects demand for trials both directly and indirectly through its influence on the number of prosecutions, i.e., caseload.[22] The theory is represented by a set of two equations—one with the prosecution rate as the dependent variable and the other with the ratio of trials to dispostions as the dependent variable. (Ordinary least squares is used to estimate the coefficients in both equations.)

The direction of causation

In both Landes's and Rhodes's models, case processing time is an independent variable affecting demand for trials. In contrast, Robert Gillespie takes case processing time as a dependent variable to be explained by demand for court services and court productivity.[23] Independent variables in Landes's and Rhodes's models are dependent variables in Gillespie's model and *vice versa.* The authors are looking at the same set of variables from slightly different points of view.

It seems reasonable to hypothesize with Rhodes that the number of defendants who demand trials depends partly upon the wait for a trial the defendant can expect and that the defendant estimates that wait on the basis of current case processing time. Rhodes justifies this ordering of case processing time and caseload on theoretical and empirical grounds.[24] He explicitly rejects a model with reciprocal causation between current demand for trials and current case processing time.

It also seems reasonable to hypothesize with Gillespie that demand for court services as reflected in current filings and pending caseloads affects current case processing time. Gillespie finds that pending caseloads (a measure of past demand), but not current filings, affect case processing time.[25] He cautions that if his assumption that demand is independent of case processing time is incorrect (the work of Rhodes and Landes would suggest that this is the case), the results of his analysis are biased.

Most models describe the court's capacity to handle cases—but not its reward structure or the motives of participants.

What is the direction of causation? Unfortunately, as I noted in discussing the general strengths and weaknesses of survey designs, we cannot empirically determine the direction of influence from one variable to another. Given the plausibility of both formulations, one might conclude that a more complete model of the relationship would specify a causal flow in both directions. In this example, the problem of unclear causal ordering could be resolved either by specifying reciprocal links between caseload and case processing time or by extending the model explicitly to include the time dimension implied by Gillespie's finding that pending caseload, but not current filings, affect case processing time and estimating the model with data collected over time. The advantage of collecting data over time is that what appear to be reciprocal links at any one time point can often be conceived of as unidirectional over time (for example, case processing time at Time 1 may affect caseload at Time 2, which in turn affects caseload at Time 3).

Although the economists who have worked on the question of court delay have not allowed reciprocal causation in their models, the requirements of specifying

22. Rhodes, *supra* n. 20, at 319–20.
23. Gillespie, *supra* n. 20, at 24, 40.
24. Rhodes, *supra* n. 20, at 324.
25. Gillespie, *supra* n. 20, at 29.

equations to represent their theories have led them to make clear and justify their choices. Possible extensions of the models grow out of the clarity with which the model must be specified. Finally, it should be emphasized that although these authors chose their variables and derived their hypotheses about the relationships among these variables from economic theory, the mathematical representation of hypotheses by means of a set of structural equations does not depend on economic theory. The hypotheses represented could be derived from economic, psychological, organizational, or any other theory of human behavior that a researcher believes explains variation in case processing time.

D. Process models

Mathematical models of court processes are the last group of models on which I want to comment. These models derive from a variety of theories, but share enough characteristics to make it useful to treat them as a class.[26] They are similar in at least three ways. First, it is usually the physical capacities of courts that are modelled. Courts are conceived of as relatively closed systems whose operations can be represented by a set of equations. The variables in these equations are such things as numbers of courtrooms, calendar size, and fraction of cases disposed. Second, the models tend to be problem oriented. The variables they include are likely to be ones over which a court decision-maker has some control. The models seek to project the consequences of alternative courses of action into the future and, given a goal or set of goals, to prescribe action for goal maximization. And, third, many of these studies use computer simulations to estimate the parameters of the models and to validate their operation.

For example, the caseflow through a court might be modelled by representing the court as a set of processing stages with queues between the stages. The court would consist of a number of courtrooms, operating for so many hours each day, and doing so many of each of a set of tasks with each task taking so many minutes. Data would be collected from the court modelled to estimate the transit times between processing points and to estimate the time needed to accomplish each processing task (e.g., time needed for a conference between a defense attorney and prosecutor or time spent by a judge waiting for the next ready case). The success of the model is judged by the similarity of its outputs to those of the court modelled. Once the model is validated, systematic alterations can be made in the model (e.g., adding another grand jury) and the effects of these alterations on the outputs of the model observed.

These models represent court structure and operations in great detail, which permits the observation of the effect of very small alterations in procedures that might otherwise go unnoticed. They also provide the ability to experiment with changes in structure or operations through simulations, which allows decisionmakers to test the effects of potential innovations inexpensively and rapidly without interrupting the work of the court. Because the variables in these models are susceptible to manipulation, the explanations they offer can be translated into action by court personnel.

Disadvantages of process models

The principal shortcomings of the models of court operations have to do with their specification. Although many of these models include great detail about the structure of the courts and the capacity of those structural elements to handle cases, other important variables and relationships are omitted. Case characteristics, the reward structure of the court, and the motives of participants are not included, but there is certainly evidence

26. For examples of these models, *see* J. Chaiken *et al.* CRIMINAL JUSTICE MODELS: AN OVERVIEW. Washington, D.C.: Government Printing Office, 1976; John B. Jennings, *Evaluation of the Manhattan Criminal Court's Master Calendar Project. Phase I—February 1, 1971–June 30, 1971.* New York: The New York City Rand Institute, 1972; John B. Jennings, "Quantitative Models of Criminal Courts," paper presented at the 29th National Meeting of the Operations Research Society of America in Dallas, Texas, 1971.

See also Jean Taylor, Joseph Navarro, and Robert Cohen. DATA ANALYSES AND SIMULATION OF THE DISTRICT OF COLUMBIA TRIAL COURT SYSTEM FOR THE PROCESSING OF FELONY DEFENDANTS. Arlington, Virginia: Institute for Defense Analyses, 1968; and Lucius J. Riccio, *A Model for Court Resource Planning,* THE JUSTICE SYSTEM JOURNAL 49 (March, 1975).

that these variables are important.

For example, a decision to reduce the length of adjournments that, according to a simulation, should result in shorter case processing time may not do so if there are participants who prefer longer case processing times and who find ways to adapt the system to their own goals. Over time, the average length of adjournment might decrease, but the number of adjournments might rise. The work of Levin, and of Banfield and Anderson, suggests that there are incentives for some court participants to prefer longer case processing times and for other participants not to oppose those who seek them. An adequate prescriptive model needs to take these variables and relationships into account.

Another specification problem in these models is that most assume that the caseload is independently produced. Again, there is evidence to the contrary. Rhodes finds that the prosecution rate depends on case processing time. The relationship between demand for court services and court productivity found by Gillespie also suggests that change in case processing time can produce changes in the relationships among the proximate variables included in the process models. There is good reason to believe that important variables have been omitted and that, over time, these variables depend on case processing time. Such relationships need to be incorporated into the models if they are to be used to test the effects of innovations.

Mathematical models of court processes can be useful decision-making and theory building tools, but the process is not independent of substance. The variance models described earlier may offer satisfying explanations, but their policy utility may be limited because many of the variables are not within the control of court actors (and, to the extent that the omitted process variables are important sources of variation, they are mis-specified). The process models, on the other hand, fail to include variables and relationships that—although outside the control of court actors—have been shown to be important so that these models are also mis-specified.

II. Developing a comprehensive theory

A comprehensive theory of case processing time would specify the relationships among all the variables believed to affect case processing time—resources, demand for court services, community characteristics, incentives, procedures, and managerial techniques. The mathematical representation of such a theory would undoubtedly require multiple equations with reciprocal links and estimation from data gathered across courts and over time. The problems of data collection, specification, and estimation inherent in doing this are enormous. We have, at this time, neither the theory nor the data to specify and estimate such a model. The goal of a comprehensive theory, however, can guide research design and data collection in more limited research projects towards those likely to yield the greatest theoretical benefits.

We can, for example, work toward a comprehensive theory by building more limited models that specify the relationships among sub-groups of interrelated variables. In this way we may be able to simplify the ultimate model building task by ruling out certain interactions or causal links. Concentration on sub-groups of variables has occurred and will continue to occur because researchers with roots in many traditions—economics, operations research, political science, and law, among others—are interested in case processing time. These differences in perspective lead to the selection of different sets of variables and relationships for analysis. The diversity of sub-theories will be most fruitful, however, if researchers are aware of the goal of a general theory and frame their questions so that links between groups of variables can be forged.

Problem selection: The goal of a comprehensive theory can also guide the selection of problems to those that seem theoretically important but have been relatively neglected in rigorous analysis. One of these is the investigation of reciprocal links and feedback loops between case processing time and variables hypothesized to affect it. There are strong reasons to expect reductions in case processing time to lead to

changes in the behavior of participants that in turn produce increases in case processing time. Reciprocal causation complicates estimation problems, but if it is not taken into account, the models may be seriously misspecified and the estimates of effects correspondingly biased.

If we extend the analysis of case processing time over time, that is, gather data on variables at a number of time points, we can simplify the problems of sorting out the causal order and estimating models. If we choose to study courts that have experienced some procedural or managerial change, study of courts over time can have additional theoretical and policy dividends. Besides making it more likely that there will be variation on the dependent variable, such courts are worth studying in order to assess the impact of particular innovations that have been suggested as a means of alleviating delay.

Data base: Data on case processing time over a sufficient period of time to estimate models with time lags are not currently available. If theories of case processing time are to be developed and estimated, however, resources will have to be invested in the accumulation of data appropriate to the theories. Constructing a data base from which the coefficients of a comprehensive model of variation in case processing time over time *and* across jurisdictions is an even longer task.To some extent, in the past, estimating techniques and data availability have guided the form of the theory rather than theory guiding the mathematical representation and choice of data. Construction of a data base will help to overcome this tendency.

A part of the problem is the development of measures of key variables that can be applied in many different jurisdictions. Both the Federal Judicial Center's District Court Studies Project and the National Center for State Court's project on court delay have addressed this problem.[27] Beyond the development of comparable measures, however, state court administrative offices and individual courts should be instructed in sampling and data collection techniques so that we can begin to accumulate the panel data—that is, data collected across cases, courts, and time—we need to estimate more complex models of case processing time.

The National Center for State Courts has prepared a set of instructions for collecting data on a sample of cases disposed in a single year.[28] Although the data gathered following their instructions can help a court estimate its case processing time, the accumulation of data suitable for estimating explanatory models of case processing time requires instructions to courts for collecting data on cases initiated at a series of time points. We know now that we need comparable data over time and across courts. The longer we postpone the design of measures and the collection of this data, the longer it will be until we are able to estimate full-fledged models of case processing time.

As this review indicates, we have the beginnings, but we do not yet have full-fledged models of case processing time. We need such models for both policy and theoretical reasons. Unless a policy-maker knows the extent to which an innovation intended to reduce case processing time can be expected to do so, he cannot reach an informed decision about the value of that innovation in relation to its cost and in relation to alternative innovations. Theoretically, too, a satisfactory explanation requires that the hypothesized causes of case processing time be specified mathematically and that the extent to which each of these variables affects case processing time be estimated. We want to know the size of the effects of variables and to be able to make causal statements about case processing time, but thus far the models, data and methods have been largely inadequate. □

27. Federal Judicial Center, DISTRICT COURT STUDIES PROJECT: INTERIM REPORT. Washington, D.C., 1976; and Thomas Church, Jr., *et al*, PRETRIAL DELAY: A REVIEW AND BIBLIOGRAPHY. Williamsburg, Virginia: National Center for State Courts in cooperation with the National Conference of Metropolitan Courts, 1978.

28. National Center for State Courts, DIAGNOSING DELAY: INSTRUCTIONS FOR CONDUCTING A TIME-LAPSE STUDY OF CASE PROCESSING TIME. San Francisco: NCSC, Reducing Trial Delay Project, 1978.

MARY LEE LUSKIN is a research associate for the American Judicature Society and associate investigator for the Evaluation of Court Delay-Reduction Programs.