COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 479

A MODEL OF THE JURY DECISION PROCESS

Alvin K. Klevorick and Michael Rothschild

January 20, 1978

A MODEL OF THE JURY DECISION PROCESS

Alvin K. Klevorick* and Michael Rothschild**

I. *Introduction*

While social scientists have had a long-standing interest in the behavior of jurors and juries, the last several years have witnessed a burgeoning of social science research on the jury. A principal stimulant of this most recent work was the set of Supreme Court decisions in the early 1970's which held constitutional the use of six-member juries *(Williams v. Florida*[1]*)* and nonunanimous twelve-member juries *(Apodaca v. Oregon,*[2] *Johnson v. Louisiana*[3]*)* in certain criminal cases in state courts and the use of six-member juries in federal civil cases *(Colgrove v. Battin*[4]*)*.

There exist several comprehensive, evaluative surveys of the jury research literature,[5] and we shall make no attempt to summarize here the broad range of studies which have been undertaken. Despite this research, however, we still know very little about how juries deliberate. As Davis, Bray, and Holt note at the conclusion of their review article, most theoretical models of the jury are relatively static in conception. Some of the recent work by Gelfand and Solomon[6] adopts a more dynamic view of jury

---

*Yale Law School and the Cowles Foundation, Yale University.

**Department of Economics, University of Wisconsin-Madison.

The authors are grateful to the National Science Foundation for financial support. We would also like to thank Don McNeil for his many contributions to the paper's development, Chris Winship for helpful discussions, and Joel Yellin for a valuable suggestion. Any errors which remain are, of course, our responsibility.

decisions. But there is still a need for theoretical models of the jury which explicitly recognize the dynamic aspects of the jury decision process. It is the goal of the present paper to develop and analyze such a model.

Dynamic models of the jury decision process are critical for a proper evaluation of the kinds of changes--to juries with fewer than twelve persons and to juries with nonunanimous decision rules--which the Court approved in the decisions mentioned earlier. The idea behind such changes is that the smaller jury or the jury requiring less of a consensus for decision will consume fewer of society's resources. Determining whether this putative economy is real can be quite complicated. For example, to assess the reduction in total trial time made possible by a change in jury size or a change in jury voting rule, the evaluator must take into account the effects these structural changes will have on the way lawyers argue the case, and evaluating such effects will be difficult. But one premise underlying suggestions for such changes is clearly the thought that smaller juries and ones which can reach verdicts based on nonunanimous majority votes will take less time to decide cases. Of course, an important question which must be asked is whether such smaller or nonunanimous juries would reach verdicts which differ from the ones their larger, unanimous counterparts would reach. If we were to view the latter juries' verdicts as the "correct" ones, then we would definitely be concerned were there to be a substantial difference between the verdicts reached by the two kinds of decisionmaking bodies. But even leaving aside the question of whether unanimous, twelve-member juries reach "correct" decisions, we are still interested in knowing whether the verdicts of the differently constituted juries will differ, how often they will differ, and what the difference will be.

In short, in evaluating proposals for changes in jury size or voting rule, one must consider the effects of such changes on resource use and on jury decisions. Given the paucity of information we have about how the jury decision process works, we have little basis for evaluating how changes in the ground rules would alter the process or even the amount of time and effort saved by such changes. The investigator who desires to gather information about the jury decision process is confronted with several different kinds of obstacles. A principal difficulty arises because it is not possible to observe the behavior of actual juries. The privacy and secrecy of actual jury deliberations must be respected.[7] As a result, the researcher is forced to turn to experiments with mock juries or to extrapolation to the jury setting of results emerging from other social psychology experiments. While mock jury experiments have grown more sophisticated and larger in scale over time, moving well beyond the anecdotal nature of earlier endeavors, there are still often significant difficulties of experimental design and inference.[8]

The belief about jury behavior which is perhaps most prevalent among those who study the jury as an institution is the "majority persuasion" hypothesis first articulated by Kalven and Zeisel in *The American Jury.* Using post-trial interviews, Kalven and Zeisel were able to reconstruct the first-ballot votes of juries in 225 criminal cases in Brooklyn and Chicago. A comparison of these first-ballot votes with the final verdicts in the cases showed that "in the instances where there is an initial majority either for conviction or for acquittal, the jury in roughly nine out of ten cases decides in the direction of the initial majority. Only with extreme infrequency does the minority succeed in persuading the majority to change its mind during the deliberation...[*W*]*ith very few exceptions the* *first ballot decides the outcome of the verdict*....The upshot is a radical hunch about the function of the deliberation process. Perhaps it does not so much decide the case as bring about the consensus, the outcome of which has been made highly likely by the distribution of first ballot votes."[9] This hypothesis, that the final verdict is largely determined by the majority of the positions jurors take *before* any deliberation occurs, has found wide support in a variety of later experimental studies.[10] It also has been used by Gelfand and Solomon in their extension and development of Poisson's models of jury verdicts.[11]

In the present paper, we provide a dynamic version of the majority persuasion hypothesis. Jurors are assumed to come into the jury room with their first-ballot, predeliberation votes in mind, and we use the concept of majority persuasion to provide a description of the way the jury then moves to a verdict. This enables us to illuminate the implications for jury verdicts of a dynamic conception of majority persuasion, as opposed to the comparative-static (first-ballot, last-ballot comparison) version of the hypothesis which has been used in previous work. In particular, the model we develop provides answers to the kinds of questions we would like to answer--questions about differences in the outcome reached and the time consumed by juries of different sizes and voting rules.

If it is correct, the model developed here has strong implications for the choice of jury institutions--for example, some might see it as adding strength to the case for adopting a 10-2 decision standard rather than a unanimous standard. The model also has the virtue of being testable without compromising the secrecy of the jury since, as we shall discuss, such a test requires essentially histogram data on the times juries take to reach decisions in a large sample of cases. The theory set forth here enables us to predict the frequency distribution of times taken to reach a verdict of guilty and the frequency distribution of times taken to reach a verdict of innocent. Hence, the minimum data requirements for each case are information on the eventual verdict and the time taken to reach the verdict. Of course, one expects that the verdict and the time required to reach it may well depend on the nature of the case, the composition of the jury trying it, and so on. Hence, the more information we have available about each case, the greater the possibility for stratification of cases, and hence the greater the potential for discerning differences in the deliberation processes of juries in different kinds of cases.

Furthermore, if the model we develop is correct, it provides researchers with a way of obtaining additional information about juries and the way they deliberate. Specifically, the model permits us to estimate the "speed of the deliberation process" (a concept to be made precise shortly) and the distribution of first-ballot votes in jury trials. The latter distribution itself constitutes exactly the kind of data needed to estimate the parameters in the model Gelfand and Solomon have developed. To date, their work on the U.S. jury system, and in particular their conclusions concerning the merits of twelve-person juries, is based upon parameters estimated from data on the 225 cases Kalven and Zeisel used as the basis for their majority persuasion "hunch." The estimated distribution of first-ballot votes which our model will yield will thus permit an indirect test of the robustness of the parameter estimates Gelfand and Solomon have derived and used in their work.

In the next section, we present the simpler form of our basic model. Then, Section III contains a discussion of some of the model's implications. The more complex version of the model is presented in Section IV together with a discussion of the way in which the model can be tested. Finally, Section V provides some concluding remarks. Ours is an abstract model and it is therefore clearly not completely true. Whether or not it is useful is an empirical--and, as we note, a decidable--issue. While we do not now have the evidence to decide this issue, it nevertheless seems important to draw out the implications of a perception about jury deliberations which has substantial currency, and to do so with a model which is testable. We now turn to that task.

II. The Discrete-Time Formulation of the Model

Our theory of the jury decision process can be formulated in two alternative ways. The simpler form, which is presented in this section, models jury deliberations as a discrete-time stochastic process. This model yields answers to questions about the probabilities with which juries subject to different voting rules will reach different verdicts. These results about decision probabilities are then used in conjunction with the more complex model, which depicts the deliberation as continuous-time stochastic process, to yield an empirically testable model of the jury decision process. We defer to Section IV our discussion of the continuous-time version.

Consider a criminal jury of N people. The jury members "hear" the case--the evidence and arguments presented by both sides, the instructions issued by the judge--and then they retire to reach a verdict. They must decide, as a group, whether the defendant is guilty or innocent of the crime of which he or she is accused; they must convict or acquit the defendant. Fix as the starting point on the time scale $(t = 0)$ the time at which the jury begins to deliberate. Let $s(t)$ denote the number of jurors who would vote for conviction at time t so that $s(t)$ is an integer in $[0,N]$. We assume that the jury continues to meet and to deliberate until it reaches a decision. If a unanimous decision is required, then the jury meets until $s(t) = 0$ or $s(t) = N$. This, of course, implies that there are no hung juries, which is counterfactual. The oft-cited statistic from *The American Jury* is that in criminal trials about 5.5% of all twelve-member juries are hung juries.[12] This is not to say that hung juries are an unimportant part of the American system of criminal justice. On the contrary, the hung jury has been called "one of the many noble features of our jury system,"[13] "a valued assurance of integrity,"[14] and "that treasured, paradoxical phenomenon."[15] But many models of jury verdicts have assumed that there are no hung juries,[16] and those models which have admitted the possibility that the jury might "hang" have been formulated in the comparative-static (first-ballot/final-verdict) mold.[17]

We propose to view $s(t)$ as a time-homogeneous Markov chain.[18] Thus, the probability distribution of $s(t+1)$ is entirely determined by $s(t)$. This has two strong implications: the jury decision process is apersonal and ahistorical. First, the evolution of the shifting balance of views on guilt and innocence does not depend on which (named) jurors would vote for one verdict or the other at any point in time. Second, given the way in which jurors would cast their ballots at time t were a vote to be taken then, the course of future votes--and, in particular, the way the jurors would divide at time $t+1$ --is independent of how long the jury has been meeting, what arguments have been made, how strong various arguments have been thought to be, how many times particular jurors have changed their minds, and so on. All that matters in determining the way a vote would come out if one were taken at time $t+1$ is the set of relationships governing the movement from one jury split to another and the way the jury would divide at time t --"the present vote."

Given the Markov assumption, the probability distribution of future votes of the jury is completely described by the $(N+1) \times (N+1)$ matrix P whose typical entry is:

$$p_{ik} \equiv \text{probability that } s(t+1) = i \text{ given that } s(t) = k .$$

That is, the transition probability p_{ik} is the probability of moving from a situation in which k people on the jury would vote for conviction to one in which i members of the jury would vote to convict the defendant. There are certain restrictions on the p_{ik}. They must all be non-negative numbers and $\sum_{i=0}^{N} p_{ik} = 1$ for $k = 0, \ldots, N$. Under a unanimous decision rule, the states 0 (acquittal verdict) and N (guilty verdict) are absorbing so that

$$(1) \qquad p_{NN} = p_{00} = 1 , \quad \text{while} \quad p_{i0} = 0 \ \ for \ \ i \neq 0 \ \ \text{and} \ \ p_{kN} = 0 \ \ \text{for} \ \ k \neq N .$$

Furthermore, to ensure that the states 0 and N are the only absorbing states, we assume that

$$(2) \qquad p_{0k}^{(N+1)} + p_{Nk}^{(N+1)} > 0 \quad \text{for} \quad k = 0, \ldots, N$$

where $P^n = (p_{ik}^{(n)})$ is the n^{th} power of the matrix P of transition probabilities. The conditions in (1) and (2) imply that acquittal $[s(t) = 0]$ and conviction $[s(t) = N]$ are the only absorbing states of the system and that it is possible to reach one of these from every other state. It is the combination of (1) and (2) which implies the counterfactual conclusion that there are no hung juries.

To see how these transition probabilities relate to quantities of interest when choosing among different jury voting rules and different jury sizes, we need to introduce two additional concepts (and, unfortunately, two more pieces of notation). First, let $q(k)$ denote the probability of an eventual conviction given that the initial vote for conviction is k. Second, let $T(k)$ denote the expected number of "ballots" required to reach a verdict given that initially k members of the jury vote for conviction. We have put "ballots" in quotation marks because we are not using $T(k)$ to describe actual balloting of the jury but rather to indicate the number of times a transition can take place in the jury's position. That is, we are using "ballot" as a mnemonic device to represent a potential shift in the number of votes for conviction (and the number of votes for acquittal). The assumed homogeneous Markov chain nature of jury deliberations then implies that

$$(3) \qquad q(k) = \sum_{j=0}^{N} p_{jk} q(j) \quad \text{for} \quad k = 1, \ldots, N\text{-}1 \quad \text{with} \quad q(0) = 0 \;, \quad q(N) = 1 \;.$$

Similarly, $T(k)$ satisfies

$$(4) \qquad T(k) = \sum_{j=0}^{N} p_{jk} T(j) + 1 \quad \text{for} \quad k = 1, \ldots, N\text{-}1 \quad \text{with} \quad T(0) = T(N) = 0 \;.$$

Given the Markov-chain assumption about jury deliberations and the conditions in equation (1) determining the first and last columns of P, a theory of the jury decision process will be completed when the remaining N-1 columns of P are specified. To this end, we make a major simplifying assumption. Specifically, we assume that one and only one juror changes his mind on the verdict at a time. This is admittedly awkward in the discrete-time model under discussion since it implies that in each time period, at each "ballot," one juror changes his position. In the continuous-time version of the model, to which we turn in Section IV, this assumption has a more palatable interpretation. (In that case, our assumption is essentially that only one juror changes his mind at any given instant and that the time between instants when jurors change their minds is a random variable with an exponential distribution.) The force of our assumption is that if the current state of the jury is k, then after its next transition, the jury will be in either state $k-1$ or state $k+1$. Hence,

$$(5) \qquad p_{k-1,k} + p_{k+1,k} = 1 \quad \text{for} \quad k = 1, \ldots, N-1 ,$$

and jury deliberations, as we have modelled them, belong to the class of birth and death processes. To complete the model of how the jury deliberates, we need specify only the $N-1$ numbers $p_{k+1,k}$ for $k = 1, \ldots, N-1$.

The particular theory we wish to consider is a stylized dynamic version of the majority persuasion hypothesis discussed in the Introduction. This "momentum" view of jury deliberations assumes that at any time the probability that a juror will switch his vote from acquittal to conviction is an increasing function of the number of jurors who voted for conviction on the previous ballot. Majority pressure builds as the deliberations progress. This is not to say that the juror who changes his vote on a particular ballot *always* moves in the direction of the majority position, but rather that a move in that direction is more likely than is a move in the opposite direction. How much more likely? That is hard to say. The simplest assumption is that majority pressure is directly proportional to the strength of the majority so that

(6) $$p_{k+1,k} = \frac{k}{N}\ ;\quad p_{k-1,k} = \frac{N-k}{N} \quad \text{for} \quad k = 1, \ldots, N-1 \ .$$

This formulation assumes that majority pressure builds at a constant rate regardless of the size of the majority. It hypothesizes that the extent of majority persuasion always reflects the relative, not the absolute, size of the majority so that, for example, the pressure exerted by a 10-out-of-12 majority would be equivalent to the pressure generated by a 5-out-of-6 majority. Some experimental evidence suggests, however, that at some point in the deliberations, the absolute size of the minority rather than its relative size matters. Researchers who hold this view would say that the lonely holdout on a six-person jury is more like the only dissenter in an 11-1 vote on a twelve-person jury than he is like a two-person minority coalition in the latter. But this issue remains an open question, and we will adopt the proportionality assumption given in (6).[19]

The transition probabilities (6) have an interpretation which relates them to more familiar classical probability models. Specifically, consider N balls allocated among two urns, one labeled C (for conviction) and the other labeled A (for acquittal). The state of the system at any time is given by the number of balls in the urn marked C. In the classical Ehrenfest model,[20] transitions in the state of the system occur when one of the N balls is chosen at random and moved from the urn it is presently in to the other urn. Thus, in the classical Ehrenfest case, the transition probabilities are

(7) $$p_{k+1,k} = \frac{N-k}{N}\ ;\quad p_{k-1,k} = \frac{k}{N} \quad \text{for all} \quad k \ .$$

In our model, transitions also occur when one of the N "balls" is chosen at random. However, instead of moving from the urn it is in to the other, the distinguished ball attracts one from the other urn to its own urn. For this reason [and because (6) is clearly a reversal of (7)], we call (6) a "reverse Ehrenfest process."

For the purpose of analyzing jury behavior, it suffices to consider the reverse Ehrenfest model in (6) as a representation of the dynamics of majority persuasion. If pressed for a probabilistic interpretation of the model, we could offer the following scenario. Imagine that after each ballot each juror makes an argument explaining and advocating his present position. Each one tries to convince the other jurors that his position is correct. Exactly one of these arguments is convincing; one juror makes a sufficiently persuasive case that he convinces one other juror to join his side. Suppose further that the juror who succeeds in convincing one of his peers is chosen at random so that the probability that any one juror will be selected to make the convincing argument is $1/N$ and is independent of the past history of deliberations. Then, the probability that a particular side in the jury's debate will gain an adherent at the next ballot is proportional to the present strength of that side, and this is exactly what (6) states. (Note how this interpretation contrasts with what we would have were the jury to follow the classical Ehrenfest structure in (7). Then, the juror who changes his mind, and hence his vote, at any particular ballot would be chosen at random.)

For the momentum or majority persuasion model described, equations (3) for the probability $q(k)$ of an eventual conviction given that the initial vote is k for conviction become:

$$(8) \qquad \begin{cases} q(k) = \frac{k}{N}q(k+1) + \left(1 - \frac{k}{N}\right)q(k-1) \quad \text{for} \quad k = 1, \ldots, N-1 \\ \\ \qquad \text{with initial conditions} \quad q(0) = 0 \, , \quad q(N) = 1 \, . \end{cases}$$

The equations (4) for the expected number of ballots required to reach a verdict given that initially k jury members vote to convict the defendant become:

$$(9) \qquad \begin{cases} T(k) = \frac{k}{N}T(k+1) + \left(1 - \frac{k}{N}\right)T(k-1) + 1 \quad \text{for} \quad k = 1, \ldots, N-1 \\ \quad \text{with initial conditions} \quad T(0) = T(N) = 0 \ . \end{cases}$$

Hence, in the reverse Ehrenfest version of majority persuasion, the eventual conviction probabilities and the expected number of ballots, each conditioned on the number of jurors who vote for conviction at the outset, are described by simple second-order linear difference equations with variable coefficients.

The reverse Ehrenfest process is a special case of birth-and-death processes studied in the literature.[21] It has some exceptionally attractive properties, including tractability, which enable us to derive analytical results about jury deliberations governed by such a process. The nonhomogeneous difference equation (9) for the expected number of ballots required to reach a verdict if the initial vote is k for conviction can only be solved using numerical methods, though it is easily solvable. On the other hand, consider the probability $q(k)$ of an eventual conviction when there are initially k votes to convict the defendant. The solution to the homogeneous second-order linear difference equation in (8) is of the form in (10)

$$(10) \qquad q(k) = A_1R_1(k) + A_2R_2(k) \ ,$$

where the $R_i(k)$ are independent solutions of (8) which we rewrite as:

$$(11) \qquad kR_i(k+1) - NR_i(k) + (N-k)R_i(k-1) = 0 \ .$$

One solution of (11), call it $R_1(k)$, is:

$$(12) \qquad R_1(k) = C$$

where C is an arbitrary constant.

To find another solution of (11), note that for any k ,

$$(13) \qquad kR_2(k) - NR_2(k) + (N-k)R_2(k) \equiv 0 \; .$$

Subtracting (13) from (11), with $i = 2$ in the latter, we find that

$$(14) \qquad k[R_2(k+1) - R_2(k)] + (N-k)[R_2(k-1) - R_2(k)] = 0 \; .$$

Let $D(k) \equiv R_2(k) - R_2(k-1)$ for $k \geqq 1$, so that

$$(15) \qquad R_2(k) = R_2(0) + \sum_{j=1}^{k} D(j) \; .$$

In this notation, (14) becomes

$$(16) \qquad D(k+1) = \frac{(N-k)}{k} D(k) \; .$$

Hence, we are seeking the solution of the first-order homogeneous equation in (16). The general solution to (16) is easily seen to be:

$$(17) \qquad D(k) = \frac{(N-(k-1))}{k-1} \cdot \frac{(N-(k-2))}{k-2} \cdot \; \ldots \; \cdot \frac{N-1}{1} D(1)$$

$$= \frac{(N-1)!}{(k-1)!([N-1]-[k-1])!} D(1) = \binom{N-1}{k-1} D(1) \; ,$$

where $\binom{n}{r}$ is the binomial coefficient. Substituting (17) into (15), we find that

$$R_2(k) = R_2(0) + \sum_{j=1}^{k} \binom{N-1}{j-1} D(1) \tag{18}$$

where $D(1)$ and $R_2(0)$ are yet to be determined.

Substituting into (10) the two independent solutions of (8) which we have found, namely, (12) and (18), we see that the solution to the difference equation for the probability $q(k)$ is:

$$q(k) = A_3 + A_4 \sum_{j=1}^{k} \binom{N-1}{j-1} \tag{19}$$

where the constants A_3 and A_4 are given by

$$A_3 = A_1C + A_2R_2(0)$$

and

$$A_4 = A_2D(1) \ .$$

The values of the constants A_3 and A_4 must satisfy the initial conditions in (8). From the first of these, $q(0) = 0$, we have $A_3 = 0$. The other initial condition, $q(N) = 1$, then implies that:

$$1 = q(N) = A_4 \sum_{j=1}^{N} \binom{N-1}{j-1} = A_4 \sum_{j=0}^{N-1} \binom{N-1}{j}$$

$$= A_4 2^{(N-1)} \ .$$

Thus $A_4 = 2^{-(N-1)}$ and $q(k)$ is given by

$$q(k) = 2^{-(N-1)} \sum_{j=0}^{k-1} \binom{N-1}{j} \ . \tag{20}$$

The expression in (20) is easily recognized as the cumulative binomial distribution with $p = 1/2$. Hence, the probability that an N-person jury which is operating under the unanimity standard and which initially records k votes for guilty eventually convicts the defendant is equal to the probability that in N-1 tosses of a fair coin, k-1 or fewer heads occur. Although the derivation of this result is straightforward, we have not been able to find a simple intuitive explanation of it.

III. Implications of the Model: Nonunanimous Decision Standards

If the reverse Ehrenfest model accurately represents the way in which juries deliberate, then we can use it to consider the effects of changes in the rules governing jury decisionmaking. Consider, for example, the effects of changing the jury decision standard to something less than a unanimity requirement. For concreteness, suppose the standard is weakened so that agreement among only 10 of the 12 jury members is sufficient for a decision. Such a change would have two effects. First, some cases would be decided differently, and second, the deliberations of any particular jury would take a different amount of time. The question is how strong are these two effects.

To approach this issue, we suppose that changing the decision standard does not alter the way the jury deliberates. This, it should be recognized, is a very strong assumption. We are supposing that despite the change in the degree of agreement required for the jury to render a verdict, the reverse Ehrenfest model continues to describe the transition probabilities of the jury's ballots. Hence, a change to a 10-2 standard leaves unchanged the transition probabilities for all states in which nine or fewer jurors agree on a position, but it changes the absorbing states from 0 and 12 to

0, 1, 2, and 10, 11, 12. We have:

$$(21)\quad \begin{cases} p_{k+1,k} = \dfrac{k}{12}\,, \quad p_{k-1,k} = \dfrac{12-k}{12} \quad \text{for} \quad k = 3, \ldots, 9\,; \\[2ex] p_{kk} = 1\,, \quad p_{ik} = 0 \quad \text{when} \quad i \neq k \quad \text{for} \quad k = 0, 1, 2, 10, 11, 12. \end{cases}$$

The difference equation for $q(k)$ under the reverse Ehrenfest assumption, which is given in (8), now applies for $3 \leqq k \leqq 9$ and the initial conditions are: $q(2) = 0$; $q(10) = 1$. The expression in (19) will provide the values for the eventual conviction probabilities, $q(k)$, for $3 \leqq k \leqq 9$ when we have found the values of the constants A_3 and A_4 which satisfy the new initial conditions.

In general, if $N-V$ is the minimum number of votes of guilty required for a conviction (so that V is the maximum number of votes for acquittal which is consistent with a verdict of guilty) and symmetrically $N-V$ is the minimum number of jurors who must vote for an acquittal for the jury to render a verdict of innocent, then:

$$A_3 = -\sum_{j=0}^{V-1}\binom{N-1}{j}\left[2^{N-1} - 2\sum_{j=0}^{V-1}\binom{N-1}{j}\right]^{-1}$$

$$A_4 = \left[2^{N-1} - 2\sum_{j=0}^{V-1}\binom{N-1}{j}\right]^{-1} .$$

In the specific case of a 10-2 rule, we find that:

$$(22)\quad \begin{cases} q(k) = -\dfrac{12}{2^{11}-24} + \dfrac{1}{2^{11}-24}\displaystyle\sum_{j=0}^{k-1}\binom{11}{j} \quad \text{for} \quad 3 \leqq k \leqq 9 \\[2ex] q(0) = q(1) = q(2) = 0\,; \quad q(10) = q(11) = q(12) = 1\,. \end{cases}$$

Table 1 presents the values of the eventual conviction probabilities conditioned on the number of jurors who vote for conviction at the outset --that is, the $q(k)$ values--for twelve-person juries which must reach a unanimous verdict and for twelve-person juries which operate subject to a 10-2 decision rule. The numbers in the column headed "Unanimity Rule" are the values for $q(k)$ in (20) with $N = 12$ while the numbers in the column headed "10-2 Standard" are calculated using equation (22). Comparing the values in the two columns, we see that given the same initial vote, juries requiring a consensus of only 10 members to reach a decision will render guilty verdicts in about the same fraction of cases as juries which must be unanimous in their ultimate decision. This comparison actually slightly understates the probability that the *same* jury would decide a specific case differently under a 12-person unanimous decision standard than under a 10-2 decision rule. The probability that the change in decision standard would lead the same jury to reach a different verdict is very small: specifically, zero if the jury is initially unanimous, .0005 if the jury is initially divided 11 to 1, and .0059 for all other initial-ballot divisions.

The explanation of this result is straightforward. If the jury initially votes 11-1 for acquittal, then under a 10-2 standard it would not deliberate further; the defendant would be found innocent. On the other hand, with unanimity required for a verdict, the same jury would continue its deliberation. But the probability that it would eventually vote to convict--that is, reverse its first-ballot majority's position--after further deliberation is, as the "Unanimity Rule" column of Table 1 indicates, only .0005. Of course, the argument is symmetric for a jury which begins with 11 out of 12 jurors voting to convict the defendant.[22] If the initial vote were 10-2 for acquittal (conviction), the jury operating subject to the 10-2 decision standard would immediately render a verdict of innocent (guilty), while the jury operating subject to a unanimous decision rule would go on discussing the case. The probability that the latter would reverse its original majority is seen from the "Unanimity Rule" column in Table 1 (specifically, the entries corresponding to 2 initial votes for conviction and 10 initial votes for conviction, respectively) to be .0059.

Suppose that the initial ballot reflects a closer division among the jurors than we have just considered; that is, suppose each coalition initially has at least 3 members. No matter which of the two jury decision rules being considered actually governs, the jury cannot render a verdict based on its members' initial votes. Consequently, under either the unanimity requirement or the weaker 10-2 decision standard, the jury will deliberate, and it will continue deliberating until it reaches a 10-2 split whether for conviction or acquittal. Then if the jury needed only a 10-2 majority for a decision, it would cease deliberating and announce its verdict. If, in contrast, unanimity was required to reach a decision, the jury would continue deliberating. The critical question is, given that the jury subject to the unanimity rule reaches a 10-2 division for conviction (acquittal), what is the likelihood that it will eventually reach the opposite verdict--that is, a unanimous position that the defendant is innocent (guilty)? Only if such a reversal occurs will the unanimity decision standard lead to a different outcome that the 10-out-of-12 rule would.
The entry in Table 1 for the probability of conviction under a 12-person unanimous standard when the initial vote is 10 for conviction, and the entry for the probability of conviction by such a jury when it begins with 2 persons voting for conviction, show that such a reversal occurs with probability

.0059. Thus, *if* the change from a unanimous decision requirement to a 10-2 majority standard does not cause the jury's deliberation pattern to depart from the reverse Ehrenfest process we have specified, that alteration in the decision standard will change decisions in less than six-tenths of one percent of the cases.

One might suspect that this conclusion about the minimal effect of the decision standard on the outcome of the jury's deliberations results because we have assumed the jury decision process is a birth-and-death process. This is not the case; the majority persuasion or momentum formulation of our model is critical to our result. One way to observe the importance of the majority-persuasion assumption is to consider an alternative birth-and-death process: specifically, a random walk. In the random-walk case, $p_{k+1,k}$ is a constant, call it γ , for all k . The probability that at any ballot the number of jurors voting to convict the defendant changes by one is not only apersonally and ahistorically determined, as in the reverse Ehrenfest model, but it is also independent of the fraction of jurors who voted to convict on the previous ballot. With this specification, the jury decision process is formally identical to the classical gambler's ruin problem. The question of whether, when operating subject to the unanimity requirement, an N-person jury which initially registers k votes to convict and $N-k$ votes to acquit eventually renders a verdict of not guilty--that is, acquits the defendant--is then analytically the same as the following question: If an individual with an initial fortune of $\$k$ continues to wager \$1 at each play of a gambling game which he wins with probability γ and loses with probability $1-\gamma$, what is the probability that he will go bankrupt before he reaches his goal of increasing his fortune to $\$N$? Our probability $q(k)$, that the jury renders a guilty verdict given that it begins with k votes for conviction is then equal to the probability that the gambler who begins with $\$k$ succeeds in attaining his goal of $\$N$ before he goes bankrupt.

The classical gambler's ruin problem has been fully analyzed in the literature, and we will present here only the results we need for our current discussion.[23] From the standard analysis, it follows that if the jury deliberations follow a random walk with the probability, at any ballot, of a transition from k votes for conviction to $k+1$ votes for conviction always equal to the constant γ (for $1 \leq k \leq N-1$) and if the jury must reach a unanimous verdict, then the probability of an eventual conviction given an initial-ballot vote of k for conviction is:

$$(23) \quad \begin{cases} q(k) = \dfrac{\left(\dfrac{1-\gamma}{\gamma}\right)^k - 1}{\left(\dfrac{1-\gamma}{\gamma}\right)^N - 1} & \text{for} \quad \gamma \neq 1/2 \\[2ex] q(k) = \dfrac{k}{N} & \gamma = 1/2 \ . \end{cases}$$

And, the expected number of "ballots" given that on the initial ballot k jurors vote to convict the defendant is

$$(24) \quad \begin{cases} T(k) = \dfrac{k}{1-2\gamma} - \dfrac{N}{1-2\gamma}\left[\dfrac{\left(\dfrac{1-\gamma}{\gamma}\right)^k - 1}{\left(\dfrac{1-\gamma}{\gamma}\right)^N - 1}\right] & \text{for} \quad \gamma \neq 1/2 \\[2ex] T(k) = k(N-k) & \text{for} \quad \gamma = 1/2 \ . \end{cases}$$

Just as we derived the result in (22) for a 10-2 standard in the majority persuasion (or reverse Ehrenfest) case, we can now compute the new values of $q(k)$ which emerge under the random walk assumption and a

10-2 verdict decision rule. We shall not go into the details of the calculations, but comparing the results of those calculations with the expression in (23) shows that under the random-walk hypothesis as many as 16.7% of the cases would be decided differently under the two verdict decision standards: unanimous 12-person verdicts and 10-out-of-12 person majority verdicts. (This largest divergence between the verdicts rendered under the two decision standards occurs when the deliberation process is modeled as a random walk with $\gamma = 1/2$.) This result contrasts sharply with the minimal effect we found the change from a unanimous to a 10-2 decision standard had on jury verdicts under our formulation of the majority-persuasion hypothesis. Hence, it should be clear that our result about how changing the decision standard in the majority persuasion model affects final verdicts derives not from the general use of birth-and-death processes but from the specific assumption about how majority pressure exerts itself in the jury room.

Turning to the question of how a change in the jury decision standard would affect deliberation time in the majority persuasion model, we must recall the artificial character of the "ballots" we have been discussing. Each ballot simply corresponds to one juror's changing his mind (and his vote), and the time between two consecutive ballots corresponds to the time between the instants at which two successive changes occur in jurors' positions. The continuous-time version of the model, to which we shall turn shortly, will provide a more plausible basis for using the number of these ballots as a measure of deliberation time. For now, it suffices to observe that if $L(k)$ is the length of time it takes to reach a verdict when there are k votes for conviction on the initial ballot, if $B(k)$ is the number of ballots required when the initial ballot is split $k/N-k$, and if $\theta(j)$

is the time which elapses between the j-1st and the j^{th} vote, then:

$$(25) \qquad L(k) = \sum_{j=1}^{B(k)} \theta(j) \ .$$

We further assume that the time between ballots--the $\theta(j)$'s --are independent random variables with a *common* distribution. In particular, we suppose that neither the length of time the jury has been deliberating nor the jury's current or previous divisions affect the probability distribution of the time between any two successive ballots. Then we can apply Wald's identity[24] to (25). Recalling that $T(k)$ is the expected number of ballots required for a verdict given an initial ballot in which k jurors vote to convict, so that $T(k) = EB(k)$, and denoting $E\theta(j)$ as $\overline{\theta}$, we have:

$$(26) \qquad EL(k) = T(k)\overline{\theta} \ .$$

Hence, with our assumptions, the effect of a change in the jury decision standard on the expected length of deliberations is directly proportional to the change's effect on the expected number of ballots--with each quantity conditioned on the number of first-ballot votes for conviction--and the proportionality factor is the expected value of the time between ballots.

Although the nonhomogeneous difference equation (9) for $T(k)$ can be solved rather easily using numberical methods, no neat analytical tricks seem to emerge to enable derivation of closed-form solutions of the type derived for $q(k)$ in (20). To determine the effect of changing from a 12-person unanimous standard to a 10-out-of-12 majority standard, we solve (9) and compare that solution with the one derived when the initial conditions in (9) are replaced by $T(2) = T(10) = 0$ and the difference equation there applies only for $3 \leqq k \leqq 9$. The results of the computations are shown in Table 2. The reduction in the expected number of ballots as one moves from a 12-person unanimity standard to a 10-out-of-12 person majority decision rule is quite striking. As the last column of Table 2 indicates, the minimum reduction in the expected number of ballots, which occurs when the jury initially splits 6-6, is 25.6%. The reduction ranges as high as 100% because when the jury begins with an 11-1 or a 10-2 split, it must deliberate to reach a unanimous decision whereas no further deliberation is required in such cases if the 10-out-of-12 decision rule is in effect.

Of course, for any initial-ballot division, whether the time saved is or is not substantial depends on the magnitude of $\bar{\theta}$, the expected time which elapses between ballots. A sharp reduction in the expected number of ballots will have only a small effect on expected deliberation time if $\bar{\theta}$ is small, while it will have a large impact if $\bar{\theta}$ is large. As we shall see, the continuous-time version of the model provides a way of estimating $\bar{\theta}$, and hence it enables us to talk more precisely about how changes in the decision standard affect expected deliberation time. We now turn to that version of the model.

IV. The Continuous-Time Formulation of the Model

Suppose the jury meets continuously. Assume, further, that the probability of one change occurring in the jury's vote division in a time period of length h is:

$$(27)\quad \begin{cases} \alpha h + o(h) \quad \text{for} \quad k = 1, \ldots, N\text{-}1 \\ \qquad 0 \quad \text{for} \quad k = 0 \quad \text{or} \quad k = N \\ \text{where } o(h) \text{ denotes a quantity of smaller order of magnitude} \\ \text{than } h \text{. That is, } o(h) \text{ is such that } h^{-1}o(h) \to 0 \text{ as } h \to 0 \text{ .} \end{cases}$$

Our majority persuasion or momentum model imposes the following further structure on the transition probabilities in (27). Given that there are currently k votes for conviction, the probability that the one change in the jury's position is to increase the number of conviction votes by one is: $\hat{\alpha}(k/N)h + o(h)$ for $k = 1, \ldots, N-1$, and the probability that the one change is to decrease the number of conviction votes by one is: $\hat{\alpha}(N-k/N)h + o(h)$ for $k = 1, \ldots, N-1$. Finally, the probability that during a time period of length h, there is *more* than one change in the jury's division of votes is $o(h)$. These assumptions comprise the characterization of a continuous-time birth-and-death process in which the probabilities of "birth" and "death" are given by our majority-persuasion specification.

Following Feller's development, we can derive the differential equations describing how the jury's position changes as it deliberates.[25] These are the differential equations for the $Q_k(t)$ variables where $Q_k(t)$ is defined as the probability that at time t there are k votes for conviction. To simplify notation, let us denote $\hat{\alpha}/N$ by α. Then, it can be shown that the differential equations governing the jury deliberation process are:

$$(28)\qquad \begin{cases} Q_k'(t) = -\alpha N Q_k(t) + \alpha(k-1)Q_{k-1}(t) + \alpha(N-k-1)Q_{k+1}(t) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for} \quad 1 \leqq k \leqq N-1 \\ Q_0'(t) = \alpha(N-1)Q_1(t) \\ Q_N'(t) = \alpha(N-1)Q_{N-1}(t) . \end{cases}$$

(Primes, as usual, denote derivatives.) The system of first-order linear differential equations in (28) can be written as:

$$
(29)\quad \begin{cases} Q'(t) = \alpha A Q(t) \\ \text{where } A \text{ is the } (N+1) \times (N+1)\text{-matrix} \\ A \equiv \begin{bmatrix} 0 & N-1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -N & N-2 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -N & N-3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & N-2 & -N & 0 \\ 0 & 0 & 0 & \cdots & \cdots & N-1 & 0 \end{bmatrix} \\ \text{and the initial conditions are } Q_I(0) = 1 \text{ , } Q_j(0) = 0 \text{ for} \\ j \neq I \text{ where } I \text{ is the number of votes to convict on the} \\ \text{first ballot.} \end{cases}
$$

The general solution to the system in (29) is of the form:

$$(30)\qquad Q(t) = e^{\alpha A t} Q(0) \text{ ,}$$

where $Q(0)$ is the vector giving the probability distribution of first-ballot votes for conviction. For example, $Q_6(0)$ is the probability that on the first ballot the jury is evenly split 6-6; and $Q_7(0)$ is the probability that initially 7 jurors vote to convict. Let $P(t) = e^{\alpha A t}$ so that $P(t)$ is an $(N+1) \times (N+1)$ matrix with the element $P_{jk}(t)$ equal to the probability that in a time period of length t a transition takes place from k jurors voting to convict *to* j jurors voting to convict. With this notation, (30) can be written as:

$$(31)\qquad Q(t) = P(t)Q(0) \text{ .}$$

It is important to note that this birth-and-death process yields the same predictions about eventual verdicts rendered as does the discrete-time reverse-Ehrenfest model we discussed earlier. That is, the probability of eventual conviction given that initially k jurors vote to convict the defendant is $P_{Nk}(\infty)$ in this model and $P_{Nk}(\infty) = q(k)$ in the reverse Ehrenfest model. To see this, note that in both models the only direct transitions from state k are to state k-1 or to state k+1 , and that the conditional probabilities of these transitions are the same in both processes. The only difference is that in the discrete-time reverse-Ehrenfest model we essentially stipulated when "ballots" were taken--when jurors changed their minds--while in the continuous-time birth-and-death process model the timing of such changes is random. But if we look at the jury deliberations only at the points at which ballots are taken and compare the transition probabilities of the two processes at those points, it is clear that the $P_{Nk}(\infty)$ probabilities will be the same as the $q(k)$ probabilities.

This identity between the predictions of eventual conviction which are generated by the two models helps us to design a test of the "majority persuasion" hypothesis as we have formulated it. Specifically, we can use the fact that $q(k) = P_{Nk}(\infty)$ to derive equations for the distribution of deliberation times for juries which reach guilty verdicts and for juries which acquit the defendant. These equations can then be used together with histogram data on such deliberation times to derive estimates of the parameters of our model. Finally, the fit of our model to such data can be compared with the fit under other hypotheses--for example, the random-walk hypothesis or a "first-ballot-determines-the-outcome" hypothesis--and we can examine how well our model predicts the distribution of times to decision on another set of data. (Of course, one way of ensuring that such another set exists is to use only part of a given data set to estimate the parameters of our model.)

To see how this can be done, let $F_N(\infty)$ be the fraction of cases in which a guilty verdict is eventually rendered. Then clearly,

$$(32) \qquad F_N(\infty) = \sum_{k=0}^{N} P_{Nk}(\infty) Q_k(0) \ .$$

Similarly, let $F_N(t)$ be the fraction of cases in which a guilty verdict is delivered in time less than or equal to t --for concreteness, say t hours. Then,

$$(33) \qquad F_N(t) = \sum_{k=0}^{N} P_{Nk}(t) Q_k(0)$$

where $Q_k(0)$ is the fraction of juries with initial ballots in which k members vote to convict (or the probability that on the initial ballot k jurors will vote to convict). Finally, define $H_N(t)$ as the fraction of guilty verdicts delivered in t or fewer hours:

$$H_N(t) = \frac{F_N(t)}{F_N(\infty)} \ ,$$

and from (32) and (33), we have:

$$(34) \qquad H_N(t) = \frac{\sum_{k=0}^{N} P_{Nk}(t) Q_k(0)}{\sum_{k=0}^{N} P_{Nk}(\infty) Q_k(0)} \ .$$

Of course, analogous quantities-- $F_0(\infty)$, $F_0(t)$, and $H_0(t)$ --can be defined for cases in which the jury eventually acquits the defendant.

The fractions $H_N(t)$ are observable. They can be easily derived from the frequency distribution of deliberation times for juries which eventually render guilty verdicts. But, as (34) suggests, our theory implies that the $H_N(t)$ values are also simple and easily calculated functions of quantities which are not observable: the $Q_k(0)$ values indicating the distribution of first-ballot votes and α which equals N times the average length of time required for a jury to change its position--that is, its vote division--by one vote.[26] Equation (34) shows clearly how $H_N(t)$ depends on $Q(0)$. It is also not difficult to see that α is the only other unknown quantity on the right-hand side of (34). First, recall that $P_{Nk}(\infty)$ equals the $q(k)$ value we calculated in the reverse Ehrenfest model. Second, $P_{Nk}(t)$ is the Nk^{th} element of the matrix $P(t) = e^{\alpha At}$ which can be calculated, with α as the only remaining parameter, because the entries of A are known and given in (29).

Hence, $H_N(t)$ can be expressed as a function of the unknown parameters[27] $Q_k(0)$ for $k = 0, 1, \ldots, N$, and α . Data on the distribution of deliberation times preceding guilty verdicts--perhaps data in histogram form--would provide observations on $H_N(t)$ for different values of t . Maximum likelihood estimation techniques--and other estimating techniques --could then be used to yield estimates of the unknowns $Q(0)$ and α . As noted earlier, we can derive an expression for $H_0(t)$ --the fraction of acquittals delivered in t or fewer hours--analogous to the expression for $H_N(t)$ in (34). Hence, the same estimation procedures could be used to derive estimates of the distribution of initial-ballot votes and the speed of the deliberation process from data on the distribution of deliberation times of juries which eventually acquit the defendant. The values of these parameters could then be compared for the two kinds of juries--

those which eventually convict and those which eventually acquit.

With the estimated values of $Q(0)$ and α in hand, the testing we described earlier could then proceed.

V. Concluding Remarks

As is the case with any abstract model, the theoretical structure we have presented and analyzed is based on a number of simplifying assumptions. We have indicated many of the model's limitations in the course of our discussion. In our ongoing work we are developing several extensions of the model presented here, extensions which will help overcome some of the current version's more striking limitations. In our view, the most important of these is the development of theoretical treatments of resistant or stubborn jurors. Our basic model makes no allowance for the possibility that some jurors are more dogged or more persuasive than others. Since anecdotal reports about jury deliberations suggest that such differences among jurors often play an important role in determining the outcome of jury deliberations, an important next step for us is to try to incorporate such differences in our model. It is to be hoped that by incorporating these differences in our model of the jury deliberation process, we can investigate theoretically (and then empirically) the importance of the effects of these juror differences. We will be able to compare the implications of the model which incorporates these differences with those of the simple model which ignores them.

Our efforts to test the model presented here and to extend it in important ways are going forward. Let us close by reiterating the point made earlier; namely, that our theory produces refutable predictions and that if the theory is true, it constitutes a powerful tool for investigating the nature of the jury decision process.

TABLE 1

EVENTUAL CONVICTION PROBABILITIES UNDER UNANIMITY AND 10-2 RULES

Initial Vote for Conviction	*Probability of Conviction*	
	Unanimity Rule	*10-2 Standard*
12	1.	1.
11	.9995	1.
10	.9941	1.
9	.9673	.9728
8	.8867	.8913
7	.7256	.7283
6	.5	.5
5	.2744	.2717
4	.1133	.1087
3	.0327	.0272
2	.0059	0
1	.0005	0
0	0	0

TABLE 2

EXPECTED NUMBER OF BALLOTS UNDER UNANIMITY AND 10-2 RULES

Initial Vote for Conviction	*Expected Number of Ballots*		
	(1) Unanimity Rule	*(2) 10-2 Standard*	$\frac{(2)}{(1)} \times 100$
12	0	0	
11	1.24	0	
10	2.89	0	
9	5.13	2.24	43.6
8	7.84	4.95	63.1
7	10.27	7.38	71.9
6	11.27	8.38	74.4
5	10.27	7.38	71.9
4	7.84	4.95	63.1
3	5.13	2.24	43.6
2	2.89	0	
1	1.24	0	
0	0	0	

APPENDIX

This appendix provides a more detailed indication of how $H_N(t)$ in (34) can be expressed as a function of the unknown parameters: $Q(0)$ and α . Since we know the values of the $P_{Nk}(\infty)$ probabilities from the reverse-Ehrenfest model calculations, we need only compute the $P_{Nk}(t)$ probabilities. The latter can be calculated using what we know about the matrix A in (29) since, by definition, $P(t) = e^{\alpha At}$.

Since A is a real, symmetric matrix, there exists an orthogonal matrix, call it X , whose columns are the eigenvectors of A , which diagonalizes A . Furthermore, the elements on the diagonal of the resulting matrix are the eigenvalues of A . That is, with the columns of X equal to the eigenvectors of A , we have:

$$\text{(A.1)} \qquad \left\{ \begin{array}{l} X^{-1}AX = D \quad \text{with} \quad D = \begin{bmatrix} \lambda_0 & & & \bigcirc \\ & \lambda_1 & & \\ & & \ddots & \\ \bigcirc & & & \lambda_N \end{bmatrix} \\ \\ \text{where } \lambda_i \text{ is the } i^{\text{th}} \text{ eigenvalue of } A . \end{array} \right.$$

But $P(t) = e^{\alpha At}$, so that $X^{-1}P(t)X = e^{\alpha X^{-1}AtX}$ and

$$P(t) = Xe^{\alpha Dt}X^{-1} \ . \quad \text{Since} \quad e^{\alpha Dt} = \begin{bmatrix} e^{\alpha\lambda_0 t} & & & 0 \\ & e^{\alpha\lambda_1 t} & & \\ & & \ddots & \\ 0 & & & e^{\alpha\lambda_N t} \end{bmatrix},$$

we have:

$$\text{(A.2)} \qquad P_{Nk}(t) = \sum_{j=0}^{N} x_{Nj} e^{\alpha\lambda_j t} x^{+}_{jk} \quad \text{for each } k \ ,$$

where x_{Nj} is the Nj^{th} element of the matrix X and x^{+}_{jk} is the jk^{th} element of the inverse matrix X^{-1} .

The expression for $P_{Nk}(t)$ in (A.2) is easily stated as a function of t with α as the only unknown parameter because the eigenvalues and eigenvectors of A are easily calculated. It can be proven, though not by elementary techniques,[28] that the eigenvalues of the $(N+1) \times (N+1)$ matrix A in (29) are: $0, -2, -4, \ldots, -2(N-1)$. There is no correspondingly simple formula for the eigenvectors of A , but they are easily calculated.

Finally, substituting into (34) the expression in (A.2) for $P_{Nk}(t)$ and the expression in (20) for $q(k) = P_{Nk}(\infty)$, we have:

$$\text{(A.3)} \qquad H_N(t) = \frac{\displaystyle\sum_{k=0}^{N} \left\{ \sum_{j=0}^{N} x_{Nj} e^{\alpha\lambda_j t} x^{+}_{jk} \right\} Q_k(0)}{\displaystyle\sum_{k=0}^{N} \left\{ 2^{-(N-1)} \sum_{j=0}^{k-1} \binom{N-1}{j} \right\} Q_k(0)} \ .$$

FOOTNOTES

[1]*Williams v. Florida* 399 U.S. 78 (1970).

[2]*Apodaca v. Oregon* 406 U.S. 404 (1972).

[3]*Johnson v. Louisiana* 406 U.S. 356 (1972).

[4]*Colgrove v. Battin* 413 U.S. 149 (1973).

[5]Davis, Bray, and Holt (forthcoming), Gelfand and Solomon (1977), Grofman (1977), Lempert (1975).

[6]Gelfand and Solomon (1977).

[7]One need only recall the problems encountered by the Chicago Jury Project when it tapped actual jury deliberations in five civil cases in the federal district court in Wichita, Kansas. Kalven and Zeisel (1966), p. xv.

[8]For excellent discussions of these problems, see Lempert (1975) and Zeisel and Diamond (1974).

[9]Kalven and Zeisel (1966), pp. 488-489 (emphasis in original).

[10]For citations to studies which provide support for the majority persuasion hypothesis, as it was put forth by Kalven and Zeisel, see Grofman (1977), pp. 5-6.

[11]See Gelfand and Solomon (1974), (1975), (1977).

[12]Kalven and Zeisel (1966), p. 56. Kalven and Zeisel also noted that in their sample the frequency of hung juries in a particular jurisdiction depended on whether that jurisdiction did or did not impose a unanimity requirement. Specifically, the frequency of hung juries in states where unanimity was required to reach a verdict was 5.6% while the frequency of hung juries in

states which allowed nonunanimous verdicts was 3.1% (p. 461). In a later article criticizing the Court's decision in *Williams v. Florida*, Zeisel (1971) presented the results of a survey of "criminal jury trials that had gone to verdict since January 1, 1969 in the Miami Circuit Court, the largest Florida court. The results were 7 hung juries in 290 trials before six-member juries," or 2.4% (p. 720).

[13]Zeisel (1972), p. 369.

[14]Kalven and Zeisel (1966), p. 453.

[15]Zeisel (1971), p. 719. As Zeisel goes on to explain in a footnote, "The hung jury is treasured because it represents the legal system's respect for the minority viewpoint that is held strongly enough to thwart the will of the majority. The paradox lies in the fact that the hung jury is only tolerable in moderation; too many hung juries would impede the effective functioning of the courts." (fn. 42, p. 719).

[16]See, for example, Walbert (1971), Gelfand and Solomon (1974), (1975), (1977).

[17]See, for example, Davis (1973), Davis, Bray and Holt (forthcoming), and Gelfand and Solomon (1975), (1977).

[18]The facts about stochastic processes which we use in this paper are entertainingly presented in Feller (Volume I, 1968), particularly, Chapters XIV and XVII.

[19]For divergent views on the issue of whether the minority faction's absolute or relative size is more important in determining the minority's resistance to majority pressure, see Zeisel (1971), (1972), and Lempert (1975)--both of whom argue that the absolute number of minority jurors is crucial--and

Grofman (1977) who argues to the contrary. Lempert's discussion also provides a good set of references to the relevant social psychology literature.

[20]See Feller (Volume I, 1968), Chapter XIV.

[21]See Karlin and McGregor (1964), (1965).

[22]As Table 1 indicates, a jury which begins with 11 out of 12 votes for conviction will render a guilty verdict with probability .9995. Hence, the probability that such a jury will, in the end, decide to acquit the defendant and thereby reverse its first-ballot majority's position is .0005.

[23]See Feller (Volume I, 1968), pp. 344-349 for a full analysis.

[24]See Feller (Volume II, 1966), Chapter XII, Section 2, Chapter XVIII, Section 2 and Breiman (1968), pp. 100-101.

[25]Feller (Volume I, 1968), pp. 454-457.

[26]Recall that $\alpha = \hat{\alpha}/N$ and that $\hat{\alpha}h$ is the (approximate) probability of one change occurring in the jury's vote division in a time period of length h .

[27]See the Appendix for a more detailed calculation of $H_N(t)$ in terms of the unknown parameters.

[28]The proof, which uses the theory of Lie groups, was discovered by Joel Yellin.

REFERENCES

Breiman, L., *Probability*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1968.

Broeder, D. W., "The University of Chicago Jury Project," 38 *Nebraska Law Review* 744 (1959).

Davis, J. H., "Group Decision and Social Interaction: A Theory of Social Decision Schemes," 80 *Psychological Review* 97 (1973).

Davis, J. H., R. M. Bray, and R. W. Holt, "The Empirical Study of Social Decision Processes in Juries" in J. L. Tapp and F. J. Levine, editors, *Law, Justice, and the Individual in Society: Psychological Legal Issues*, Holt, Rinehart, and Winston, Inc., forthcoming.

Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley and Sons, Inc., New York, Volume I (Third Edition, 1968), Volume II (First Edition, 1966).

Gelfand, A. E. and H. Solomon, "Analyzing the Decision-Making Process of the American Jury," 70 *Journal of the American Statistical Association* 305 (1975).

Gelfand, A. E. and H. Solomon, "Considerations in Building Jury Behavior Models and in Comparing Jury Schemes: An Argument in Favor of Twelve Member Juries," 17 *Jurimetrics Journal* 292 (1977).

Gelfand, A. E. and H. Solomon, "Modeling Jury Verdicts in the American Legal System," 69 *Journal of the American Statistical Association* 32 (1974).

Gelfand, A. E. and H. Solomon, "A Study of Poisson's Models for Jury Verdicts in Criminal and Civil Trials," 68 *Journal of the American Statistical Association* 271 (1973).

Grofman, B., "Theory and Experiment on Juries: Did the Supreme Court Get It Right?" Social Science Working Paper 124, School of Social Sciences, University of California, Irvine, June 1977.

Kalven, H., Jr. and H. Zeisel, *The American Jury*, Little, Brown & Co., Boston, 1966.

Karlin, S., and J. McGregor, "Ehrenfest Urn Models," 2 *Journal of Applied Probability* 352 (1965).

Karlin, S. and J. McGregor, "On Some Stochastic Models in Genetics" in J. Gurlund, editor, *Stochastic Models in Medicine and Biology*, Army Mathematics Research Center, University of Wisconsin Press, Madison, 1964.

Lempert, R. O., "Uncovering 'Nondiscernible' Differences: Empirical Research and the Jury-Size Cases," 73 *Michigan Law Review* 643 (1975).

Walbert, D. F., "The Effect of Jury Size on the Probability of Conviction: An Evaluation of Williams v. Florida," 22 *Case Western Reserve Law Review* 529 (1971).

Zeisel, H., "...And Then There Were None: The Diminution of the Federal Jury," 38 *University of Chicago Law Review* 710 (1971).

Zeisel, H., "The Waning of the American Jury," 58 *American Bar Association Journal* 367 (1972).

Zeisel, H. and S. S. Diamond, "'Convincing Empirical Evidence' on the Six-Member Jury," 41 *University of Chicago Law Review* 281 (1974).