COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 254

STOCHASTIC THEORIES OF PARTICIPATION

Joseph B. Kadane, Gordon H. Lewis, and John G. Ramage

August 19, 1968

ABSTRACT

A recent mathematical model by Horvath which predicts the process of participation in group discussions is compared to several alternative models. The models can be applied to data for which the basic unit is the act and to data for which the basic unit is a continuous sequence of acts. In both cases, when statistical tests of significance are employed, most of the more restrictive models are rejected.

STOCHASTIC THEORIES OF PARTICIPATION

by

Joseph B. Kadane, Gordon H. Lewis, and John G. Ramage*

Introduction

Interpersonal interaction, especially as displayed in group discussions, has been a long standing, if somewhat sporadic, interest of those concerned with social behavior. Although the data which have been obtained are typically of the type which lends itself to the construction of formal theories, few process theories have emerged for other than the dyadic case. A notable exception is a model proposed by William Horvath. The present paper is concerned with the adequacy of that model and several alternatives to it.

Background

In 1951 Bales published a method for studying group discussion which had as its observational unit "the smallest segment of behavior that had meaning to the group." This unit might be either

---

*The authors wish to thank Philip Bonacich, Michael Farrell, Theodore Mills, Zvi Namenwirth, Mrs. Michael Olmsted, and Frederick Strodtbeck for providing data upon which analyses in this study were based. Funds for computation of the analyses were provided through the Yale Computer Center under National Science Foundation Grant NSF GP-4774 for which the authors also express their appreciation. This research was partially supported by a grant from the National Science Foundation.

verbal or nonverbal, but it was usually a simple subject-predicate combination. Subsequently, Bales, Strodtbeck, Mills, and Roseborough (1951) published a paper in which they reported regularities observed in their studies using Bales' observational method. Of specific importance was the regularity which they felt existed in the distribution of participation, when the members of a group were ordered according to the number of acts they produced and when groups of the same size were aggregated. They suggested that the harmonic function might describe the regularity, but they rejected this notion after using a χ^2 test of goodness of fit on data from groups of size six, the groups for which the harmonic function appeared to have the best fit.

In 1952 Stephan and Mishler reported on data gathered at Princeton. In this instance the unit of observation was a "participation" which consisted of

> the word, sentence, or longer statement of an individual that follows such a participation by one member and continues until it is terminated by an appreciable pause or by the participation of another member (1952: 482).

The unit for Stephan and Mishler bears a close relation to the unit of Bales: except for the inclusion of the condition dealing with "appreciable pauses" and the apparent exclusion of non-verbal acts, a "participation" would be a complete sequence of acts by a given actor. Because of the difference between the two units, however, we shall distinguish between a "participation," and the extended sequence of acts which we shall call a "burst."

Stephan and Mishler also analyzed their data by first ordering the members of the group from highest to lowest, according to the number of times each person participated, and then aggregating the ordered data for groups of a given size. Stephan and Mishler suggested that the resulting distribution could be described by an exponential function

$$p_i = ar^{i-1} \tag{1}$$

in which p_i is the percentage for persons of rank i, and a and r are parameters depending on the size of the group. On the basis of the typically small differences between the observed and the estimated distributions they concluded that the exponential function gave a "good fit."

Stephan (1952) also attempted to evaluate the extent to which the exponential might describe the regularity in Bales' data. Although he had only one frequency distribution to work with, that for groups of size six, and although he omitted any predictions for the highest two participants, he concluded that the exponential function did indeed fit quite well. Again, the criterion for goodness of fit was not a statistical test, but the amount of error between the observed and the estimated values.

In 1965 Horvath proposed a formal theory with two assumptions:

Assumption 1: In an n person group in which there exists a hierarchy, the opportunity to speak goes to person i before person $i+1$, and either if person i speaks or if $i = n+1$, the opportunity returns again to person 1.

Assumption 2: If person i is given an opportunity to speak, he will speak with probability ϕ and not speak with probility $1-\phi$.

Horvath showed that these two assumptions imply that the probability that the i^{th} person in the hierarchy will speak next is

$$(2) \qquad \frac{\phi(1-\phi)^{i-1}}{1-(1-\phi)^{n}} .$$

Since the process described by the assumptions is one of independent trials, the proportion of acts that a person will make in the long run has the same value as the probability that he will make the next act. Hence, substituting $\phi/(1-(1-\phi)^n)$ for $\underline{a}$ and $1-\phi$ for $\underline{r}$, the result in (2) is equivalent to the Stephan and Mishler equation.

In discussing the validity of the assumptions, Horvath concluded that it was intuitively reasonably to assume that actors are characterized by the same probability value, ϕ, and that the assumption of the existence of a hierarchy, while difficult to confirm experimentally, could be given a plausible interpretation. He did not discuss any tests of his model. Perhaps, because his model predicted that Stephan and Mishler equation, and that equation was believed to describe the existing data well, the model did not appear to need further testing. But the model goes beyond predicting merely the overall proportion of acts initiated in the group; as a process model it is possible to derive many quantities, such as the probability of k consecutive acts by a person of rank j, the average number of consecutive acts by a person of rank j, and so forth. The fact that the model is in accord with a known (or presumed) regularity is a necessary but not a sufficient condition for not rejecting the model. One is left, then, with an apparent regularity and with a theory which has as one of its implications the prediction of that regularity. The unanswered question is how well the model describes the process of participation.

The Present Research

The evaluation of models and hypotheses has been and remains an area of controversy. In evaluating the goodness of fit of the harmonic function, Bales et al., used a chi-square test; in evaluating the goodness of fit of the exponential function, Stephan and Mishler presented data on the errors between the observed and the estimated values. The difference in approach is indicative of differences among researchers on the evaluation of goodness of fit. In commenting on the problem of choosing a criterion for evaluating goodness of fit, Coleman (1960:50) has pointed out that

> A statistical test like χ^2 might be used for such a criterion, but a rejection of the model by this (or another) statistical test might easily occur if the amount of data was large even if the model fit it very well.

In the present instance the amount of data available is at least as great as that which existed when Coleman made his comment, and the authors are very much aware of the significance of the point he raised. Nevertheless, we have decided, as a first step, to evaluate Horvath's model by means of statistical tests of significance.

In the following analyses, we have considered a number of different models in addition to Horvath's model, and we have tested various alterations by means of likelihood ratio and chi-square tests as suggested by Anderson and Goodman (1957). Only the likelihood ratio test will be reported since the results were substantially the same for the different tests, and the conclusions one would reach did not differ.

Maximum likelihood estimates have been used in estimating the parameters for the tests. Finding the maximum likelihood estimate, ϕ , for the Horvath models requires determining a root of an $(n-1)$ degree polynomial for a group of size n .[1] For the multinomial models, the relevant proportions are used in each case. The use of maximum likelihood estimates requires a stochastic model which determines the distribution of the observations.

Data

Four data sets were made available to the authors. Zvi Namenwirth and Michael Farrell provided data from 15 four person discussion groups which met in 1964. Each group consisted of previously unacquainted undergraduates at Yale who met for 20 minutes to discuss a "human relations" problem. The discussion was tape recorded and scored from the tape with the aid of notes made by observers. The data in their final form consisted of the total number of acts initiated by each person during the group meeting.

Mrs. Michael Olmsted and Ted Mills provided data gathered at Harvard by Michael Olmsted (1954) on 24 person groups. These data, to be referred to as the Harvard I data set, were also summary frequencies of acts initiated during the entire group discussion.

Data from 67 three person family groups in Chicago were provided by Fred Strodtbeck and Margaret Packman Ray.[2] Each group consisted of a father, mother, and a high school age son or daughter. Sequential act data were provided, but because of the special manner in which the discussions were conducted only the summary frequencies have been used in the present report.

Finally, Phillip Bonacich (1968) provided original interaction scrolls from 28 five person groups which met at Harvard. The data from these scrolls were transcribed and keypunched at Yale. The data came from 10 groups of male participants, Harvard II-M, and 18 groups of female participants, Harvard II-F. Each of the groups met for a two hour period and was composed of strangers. The two hour period was divided into four sessions: the first and third were task oriented discussion sessions, the second and fourth were non-task oriented discussion sessions. These data are, of course, sequential in nature. Sequential data are especially valuable because they can be used to test both act and burst models and because they permit tests of the assumptions about the nature of the process.

Since the Horvath model lends itself most easily to an act interpretation we shall consider the acts models first.

Acts Models

Horvath's model assumes that in a given group every person is characterized by a common value of ϕ. Because of apparent differences in the distribution of participation across groups, we have chosen to consider two variants of the Horvath model: one in which there is but a single value of ϕ which is common not only to the members of a given group but to members of all groups of the type under consideration, and one in which the value of ϕ may vary from group to group. The former shall be referred to as H_S and the latter as H_D. That ϕ _may_ vary from group to group does not imply that it does, and thus H_S is a special case of H_D. Using the symbol for inclusion, $\subset$, we shall denote the relationship between these models by $H_S \subset H_D$.

The second two models to be considered are multinomial generalizations of H_S and H_D. In the Horvath models the probability of an act by a given person is a function of the probability of an act by any other person. More general models might require only that the set of values, $\{p_i\}$, sums to unity. This could be true where $\{p_i\}$ is constant over groups, model M_S , or where it may depend on the group, model M_D . Any data set which conformed to Horvath's model would also satisfy the corresponding multinomial model, but not vice versa. The relations among these models are as follows:

$$\begin{array}{ccccc} & & M_S & & \\ & \subset & & \subset & \\ H_S & & & & M_D \\ & \subset & & \subset & \\ & & H_D & & \end{array}$$

Before turning to likelihood ratio tests it is instructive to look at a two dimensional picture of the parameter space of the Horvath models and the multinomial models. Each of the models determines a set of probabilities, $\{p_i\}$, that person i will be the next to speak. For the multinomial models the maximum likelihood estimate of p_i is $\hat{p}_i$, where $\hat{p}_i$ refers to the proportion of acts by the person of rank i within a single group for model M_D and to the proportion obtained from the ordered, aggregated frequency distribution for model M_S . Since the data are always ordered so that the person who initiated the i^{th} greatest number of acts is called person i , if one plots $\hat{p}_1$ and $\hat{p}_2$, the proportions of acts by the first and second ranked group members, the parameter space for M_D is the triangle determined by the constraints

1) $p_1 \geq p_2 \geq \cdots \geq p_n \geq 0$

2) $\sum_{i=1}^{n} p_i = 1$

or alternatively, by the equations

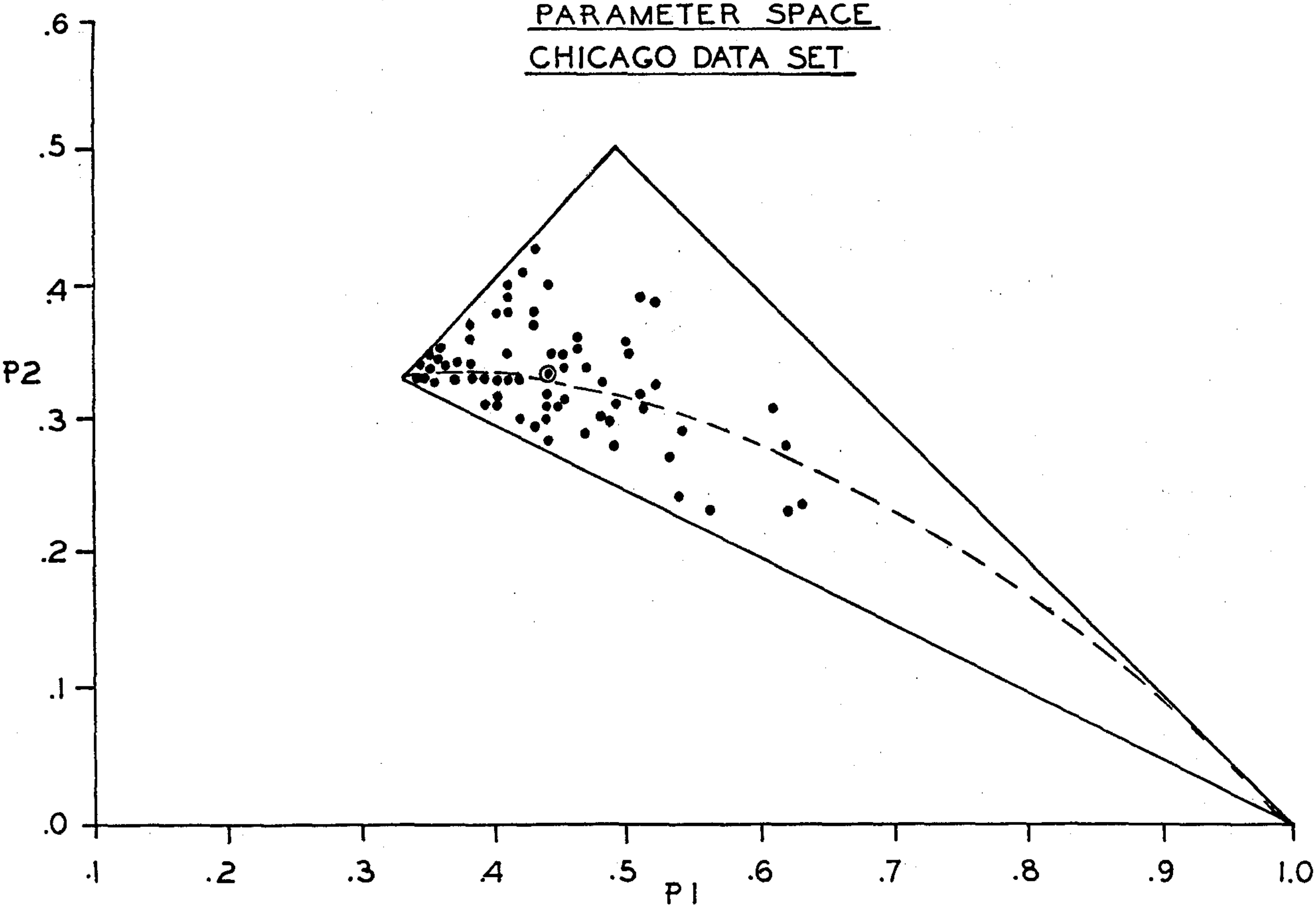
1) $p_1 + p_2 = 1.0$

2) $p_1 = p_2$

3) $p_1 + (n-1)p_2 = 1.0$.

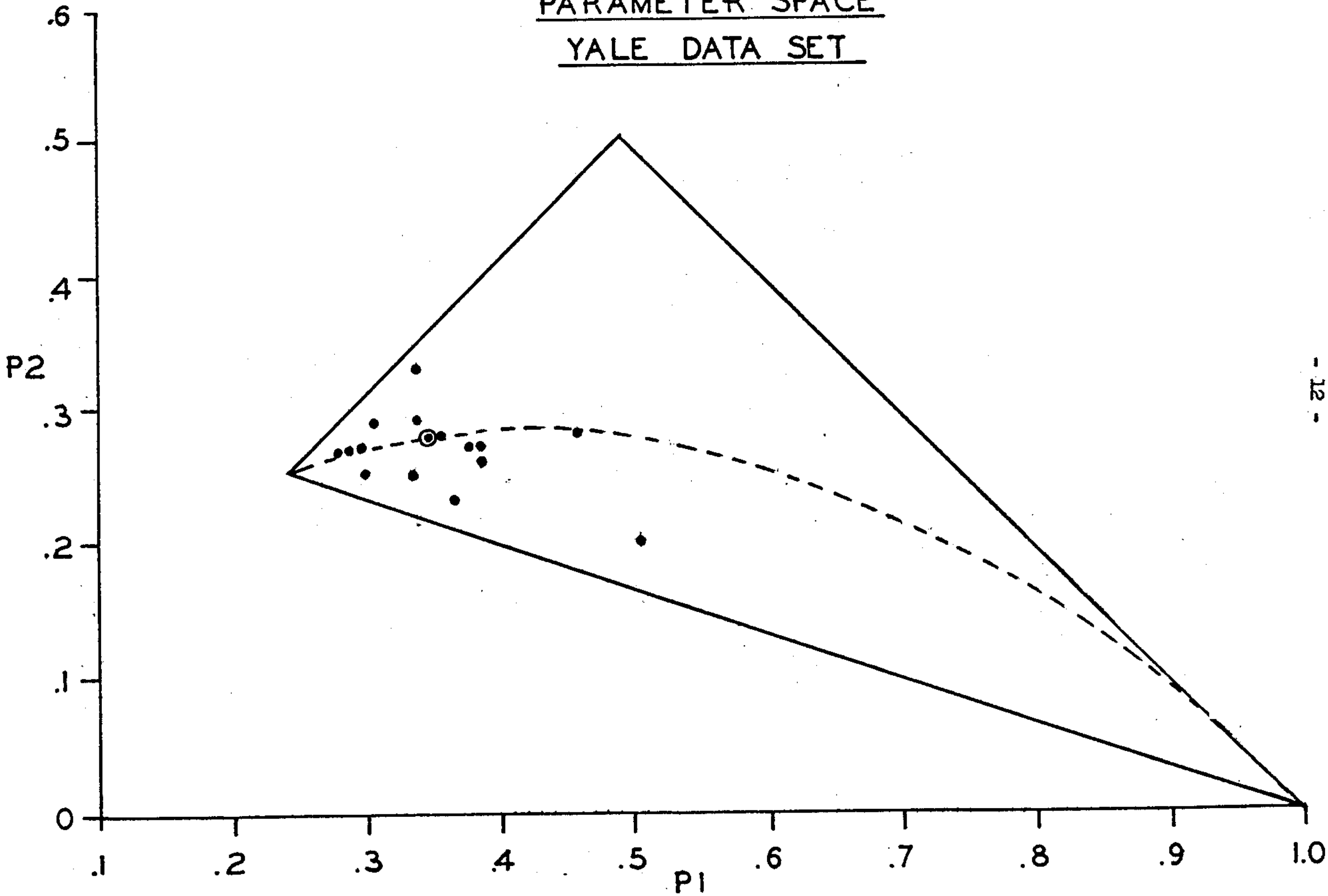
For M_S the space is reduced to a single point anywhere within this region, for H_D it is the curved line within the region, and for H_S it is any single point on the curved line. If the points $(\hat{p}_1, \hat{p}_2)$ were (a) scattered throughout the entire region, (b) concentrated at a point not on the Horvath line, (c) located along the Horvath line, or (d) concentrated at a single point on the Horvath line, it would be evidence of in favor of models M_D , M_S , H_D , or H_S , respectively. Graphs 1 through 2 show the actual distribution of values $(\hat{p}_1, \hat{p}_2)$ for the data from Chicago, Yale, Harvard I, Harvard II-M, and Harvard II-F.

In each of the graphs the points are fairly well scattered, even though points do not tend to appear in the lower right and the upper vertices. Points in the lower right vertex or the upper vertex would indicate domination of the conversation by the highest participant



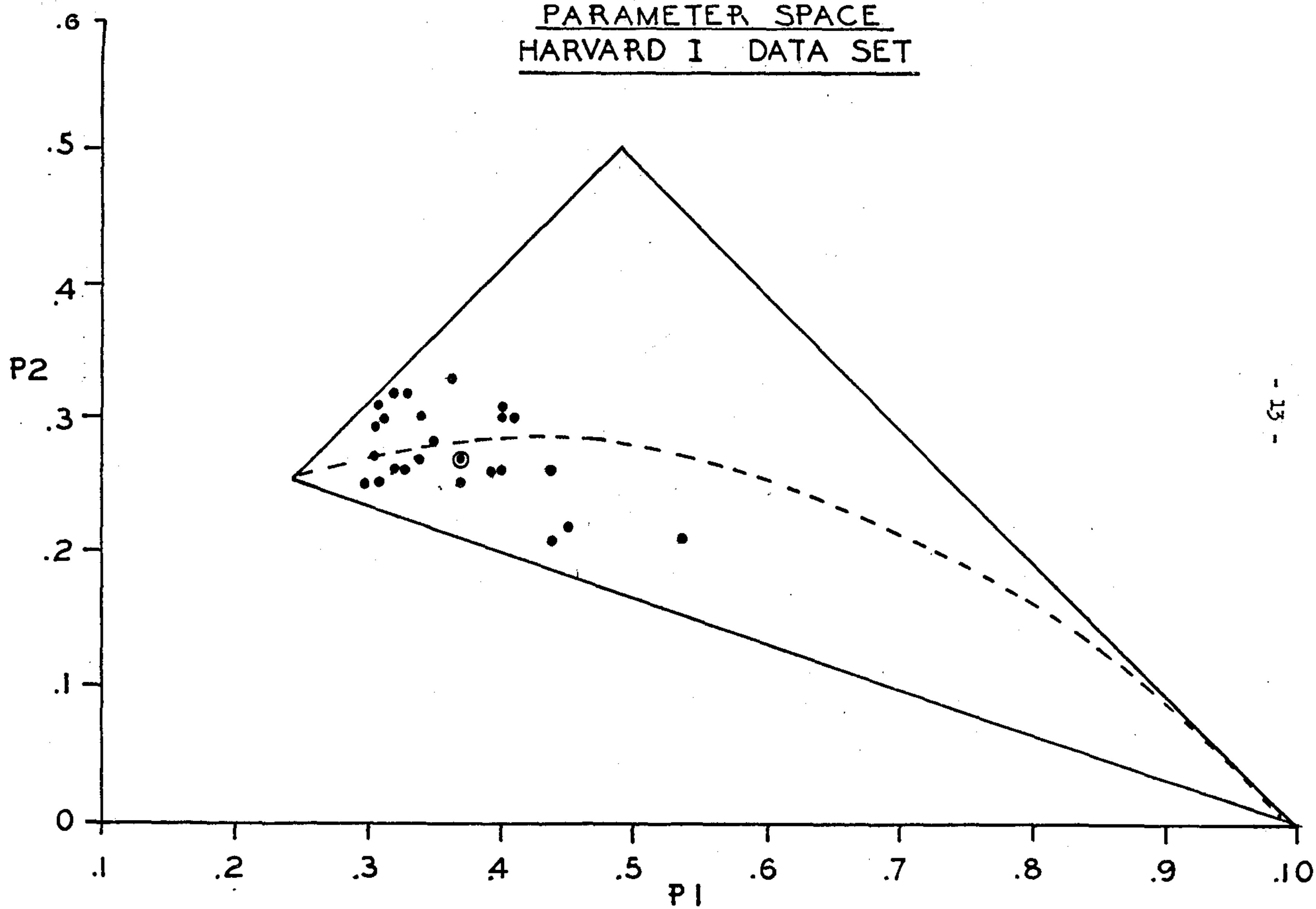
GRAPH 1
PARAMETER SPACE
CHICAGO DATA SET

GRAPH 2
PARAMETER SPACE
YALE DATA SET



GRAPH 3

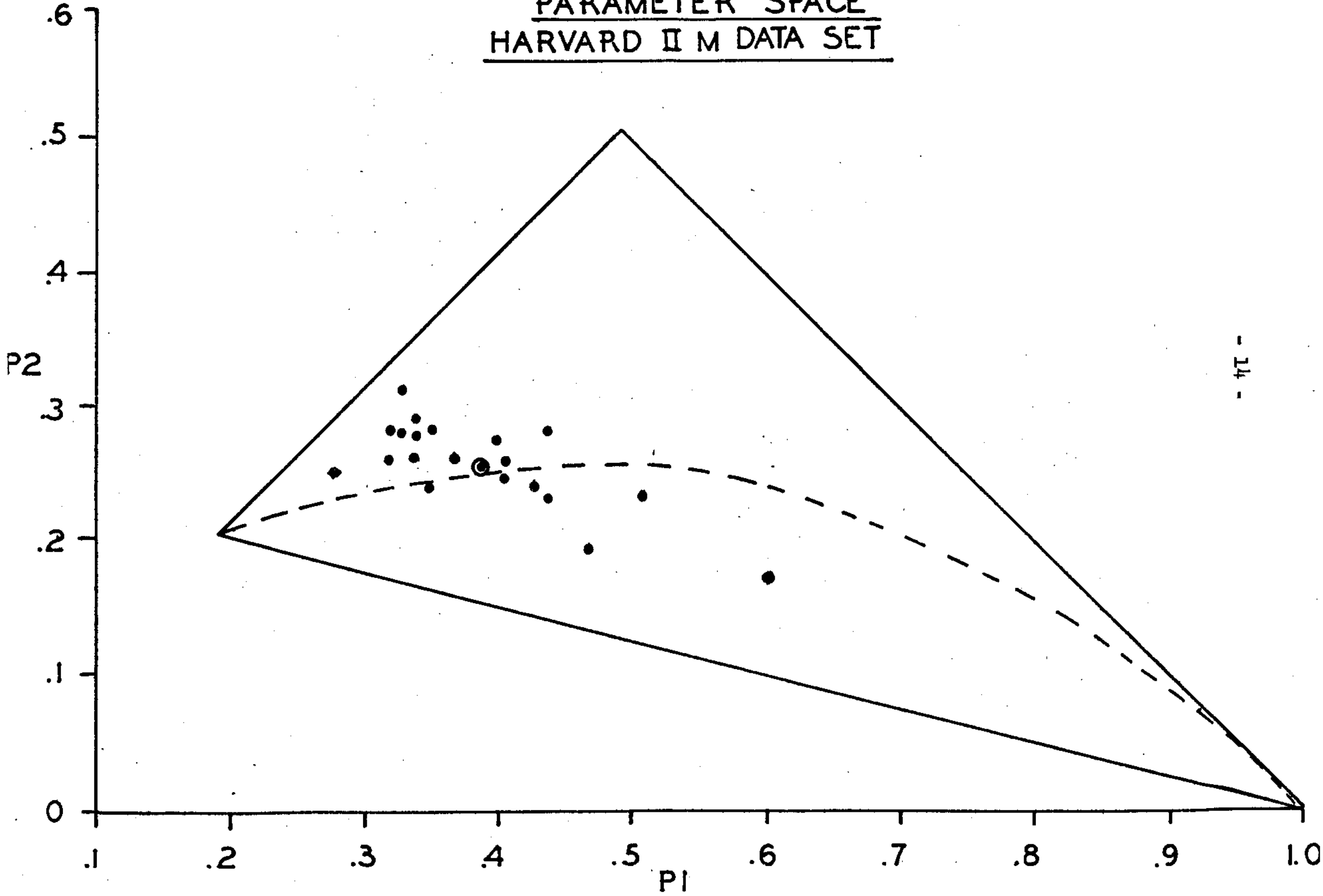
PARAMETER SPACE
HARVARD I DATA SET



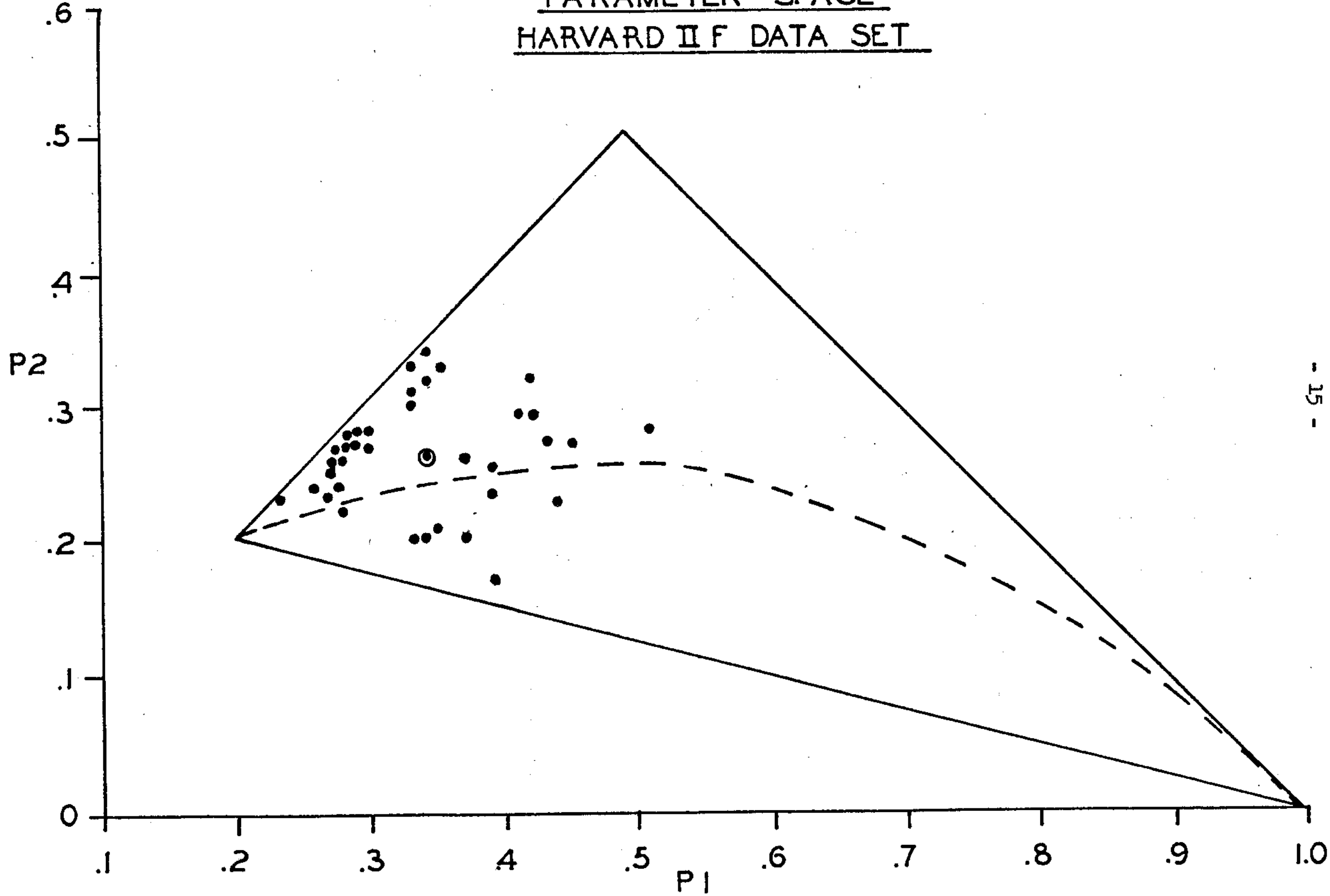
GRAPH 4

PARAMETER SPACE

HARVARD II M DATA SET



GRAPH 5
PARAMETER SPACE
HARVARD II F DATA SET



or the two highest participants, respectively. The data from groups composed only of women, Harvard II-F, Graph 5, show an interesting concentration near the line $p_1 = p_2$, which indicates a greater equality of participation between the top two actors in that data set than in any of the other data sets. In each of the graphs, however, the dispersion of points is evidence for model M_D.

If one aggregates the number of acts by persons of the same rank order in acts initiated and then proportionalizes, the points $(\hat{\underline{p}}_1, \hat{\underline{p}}_2)$, representing the aggregate proportions $\hat{\underline{p}}_1$ and $\hat{\underline{p}}_2$, are in most cases very close to the Horvath line. These points are enclosed on the graphs by the small circles.

To supplement a subjective reaction to the graphs of p_1 and p_2, likelihood ratio tests of goodness of fit were performed.[3] The results of these tests are contained in Table 1. The Horvath model with the same parameter set does poorly against the Horvath model with different parameter sets; the multinomial model with a single parameter set does poorly against the multinomial model with different parameter sets. In fact, in four of the five tests indicated in Table 1 the more restrictive model does very poorly when compared to the less restrictive model. The only test in which this is not entirely so is the test of the Horvath model with a single parameter for all groups against the multinomial model with single parameter for all groups. Even there all but one of the results are at levels which would ordinarily be associated with rejection of the hypothesis.

TABLE 1

Results of Tests of Zero Order
Horvath Multinomial Act Models

Hypothesis	Data Set	LRT	DF	≅ Sig.
H_S vs. H_D	Chicago	908.37	66	$< 10^{-6}$
	Yale	238.81	14	$< 10^{-6}$
	Harvard I	230.34	23	$< 10^{-6}$
	Harvard II-M	320.32	19	$< 10^{-6}$
	Harvard II-F	854.35	35	$< 10^{-6}$
H_S vs. M_S	Chicago	10.15	1	.001
	Yale	3.19	2	.20
	Harvard I	16.68	2	.0002
	Harvard II-M	13.98	3	.003
	Harvard II-F	270.22	3	$< 10^{-6}$
H_S vs. M_D	Chicago	1201.82	133	$< 10^{-6}$
	Yale	369.04	44	$< 10^{-6}$
	Harvard I	463.44	71	$< 10^{-6}$
	Harvard II-M	575.51	79	$< 10^{-6}$
	Harvard II-F	2015.33	143	$< 10^{-6}$
H_D vs. M_D	Chicago	293.46	67	$< 10^{-6}$
	Yale	130.23	30	$< 10^{-6}$
	Harvard I	233.10	48	$< 10^{-6}$
	Harvard II-M	255.19	60	$< 10^{-6}$
	Harvard II-F	1160.99	108	$< 10^{-6}$
M_S vs. M_D	Chicago	1191.67	132	$< 10^{-6}$
	Yale	365.85	42	$< 10^{-6}$
	Harvard I	446.76	69	$< 10^{-6}$
	Harvard II-M	561.53	76	$< 10^{-6}$
	Harvard II-F	1745.12	140	$< 10^{-6}$

There is, however, an interesting result which is masked in lumping together the results from all of the groups in each data set. As is shown in Table 1, if one tests the Horvath model with different parameters for each group against the multinomial model with different parameters for each group, H_D vs. M_D, and aggregates the statistics, the combined result significantly favors the multinomial model. If one contrasts these two models for individual groups, he finds that for part of the groups the Horvath model does quite well. Table 2 shows a typical set of results; in 9 of the 15 Yale groups the probability that the test statistic would be this large or larger is

TABLE 2

Testing H_D vs. M_D, Group by Group
Yale Data, Four-Man Groups, DF=2

Group	LRT	Sig.
1	4.08	.13
2	1.42	.50
3	3.94	.14
4	20.11	$< 10^{-4}$
5	10.70	.005
6	25.48	$< 10^{-5}$
7	18.36	10^{-4}
8	.90	.64
9	3.03	.22
10	21.62	$< 10^{-4}$
11	16.48	.0003
12	.15	.93
13	2.81	.25
14	.63	.73
15	.50	.78

greater than .05, if H_D is true. This pattern, which is similar in the other data sets is summarized in Table 3, for three different significance levels. For the Chicago, Yale, Harvard I, and Harvard II-M data sets, the Horvath distribution is not rejected at the .05

TABLE 3

<u>Proportion of Groups in which H_D vs. M_D Yielded a Significance Value $\geq$ P</u>

Data Set	Number of Groups	P .05	.01	.001
Chicago	67	.73	.76	.87
Yale	15	.60	.60	.67
Harvard I	24	.46	.63	.71
Harvard II-M	20	.45	.60	.75
Harvard II-F	36	.22	.25	.36

level in 50% of the groups, it is not rejected at the .01 level in 60% of the groups, and it is not rejected at the .001 level in 70% of the groups. This is not the pattern for the Harvard II-F set.

Previously it was mentioned that a fundamental assumption of the Horvath model was the existence of a hierarchy which determines the distribution of opportunities to act. We have assumed that in the Yale and Harvard groups under study this hierarchy arises through the interaction process. In the Harvard II-M and Harvard II-F groups which held two task oriented discussions, one would expect the second task oriented discussion to be better characterized by the Horvath model. In fact, when one tests H_D vs. M_D, more than half of the groups (17 of 28) showed a larger likelihood ratio value for the second task session (session three) than for the first. This outcome characterized both male groups (7 of 10) and female groups (10 of 18), and weakly suggests that the fit to a Horvathian process is no better in session three than in session one.

To this point then it can be said, but not unequivocally, that the Horvath model using a single parameter across groups compares unfavorably to the multinomial model with a single parameter set across groups, that for many groups the Horvath model does fairly well against the multinomial model, and that the process does not seem to be any better described in session three than in session one.

So far, all of our tests have assumed that the process is one of independent trials. Horvath's theory and its multinomial generalizations have assumed that knowing who the present actor is does not change our predictions about who the next actor will be. It seems reasonable, however, to think that conversations might be characterized by dependence from act to act, rather than by independence; it would seem that knowing who initiated the last act would make a difference in who initiates the next act. The sequential data from the Harvard II data sets make it possible for us to investigate whether the acts are independent, or if not, whether they are dependent in certain ways.

Independent trials processes, a special case of Markov processes, are sometimes referred to as Markov chains of order zero. In general, a Markov chain of order i is one in which the probability that the next state ("actor") is state j depends at most on the i previous states. The models we have investigated to this point have been Markov chains of order zero.

The main problem in using higher order Markov chains is the number of parameters requires. For an unrestricted n state Markov chain, the i^{th} order process has $n^i(n-1)$ independent parameters. An unrestricted third order chain for a five person group, for example, already requires 500 parameters. Since limitations on the duration of a conversation restrict the length of the basic acts sequence, it would be foolish to try to fit a model with 500 parameters to a sequence of 800 observed acts. For this reason consideration is limited to a hierarchy of zero, first, and second order models.

Previously the multinomial model which allowed a different set of parameters for each group was denoted M_D . Here it is necessary to mention explicitly the order of the process, so model M_D is now denoted $M_{D,0}$. Similarly, $M_{D,1}$ and $M_{D,2}$ are respectively the first and second order Markov models in which parameters may differ across groups. One can also conceive of comparable models in which it is assumed that there is a single parameter set for all groups of a given size. These models are designated $M_{S,0}$, $M_{S,1}$, and $M_{S,2}$. The relation of these models follows:

$$\begin{array}{ccccccc} M_D & = & M_{D,0} & \subset & M_{D,1} & \subset & M_{D,2} \\ & & \cup & & \cup & & \cup \\ M_S & = & M_{S,0} & \subset & M_{S,1} & \subset & M_{S,2} \end{array}$$

This hierarchy of models allows two types of tests: within a row we are testing how well the lower order processes compare to the higher order processes; within a column we are testing how well a parameter set for all groups compares to a parameter set for each group.

The results of these tests are to be found in Table 4. The

TABLE 4

Results of Tests of Multinomial Acts Models of Orders 0, 1, and 2

Hypothesis	Data Set	LRT	DF	≃ Sig.
$M_{S,0}$ vs. $M_{S,1}$	Harvard II-M	468.01	16	$< 10^{-6}$
	Harvard II-F	737.86	16	$< 10^{-6}$
$M_{S,0}$ vs. $M_{D,0}$	Harvard II-M	560.10	76	$< 10^{-6}$
	Harvard II-F	1745.12	140	$< 10^{-6}$
$M_{D,0}$ vs. $M_{D,1}$	Harvard II-M	1120.23	320	$< 10^{-6}$
	Harvard II-F	1595.97	576	$< 10^{-6}$
$M_{S,1}$ vs. $M_{D,1}$	Harvard II-M	1211.08	380	$< 10^{-6}$
	Harvard II-F	2596.60	700	$< 10^{-6}$
$M_{S,1}$ vs. $M_{S,2}$	Harvard II-M	724.05	80	$< 10^{-6}$
	Harvard II-F	809.15	80	$< 10^{-6}$
$M_{D,1}$ vs. $M_{D,2}$	Harvard II-M	2417.91	1600	$< 10^{-6}$
	Harvard II-F	3935.27	2880	$< 10^{-6}$
$M_{S,2}$ vs. $M_{D,2}$	Harvard II-M	2901.85	1900	$< 10^{-6}$
	Harvard II-F	5715.92	3500	$< 10^{-6}$

summary tests all reject the more restrictive models; even when one looks group by group at the test of zero versus first order and first versus second order, the results are still strongly in favor of the higher order process. Table 5 reports the proportion of groups in which the lower order model was not rejected at three significance levels. At the .001 level, in 10 of the 20 groups of males and 15 of the 36 groups of females the zero order representation was not rejected when compared to the first order representation. At the .001 level, in 11 of the 20 groups of males and 27 of the 36 groups of females the first order representation was not rejected when compared to the second.

TABLE 5

Proportion of Groups in which $M_{D,0}$ vs. $M_{D,1}$ and $M_{D,1}$ vs. $M_{D,2}$ Yielded a Significance Value $\geq$ P

Hypothesis	Data Set	N	P .05	P .01	P .001
$M_{D,0}$ vs. $M_{D,1}$	Harvard II-M	20	.20	.35	.50
	Harvard II-F	36	.17	.28	.42
$M_{D,1}$ vs. $M_{D,2}$	Harvard II-M	20	.30	.45	.55
	Harvard II-F	36	.42	.58	.75

In summary, although the Horvath models appear fairly good in some respects, the assumption of indpendent trials seems unsupportable. The data from some groups of females are fairly well described by a first order Markov chain, but the data from fewer groups of males are. To know whether higher order chains will more completely describe these processes requires further data.

It is also possible, however, that sequences of acts will not turn out to be easily describable by simple models. Certainly there are a number of problems in using acts as units of observation. Unreliability in categorizing acts, although of less significance in the present case than unreliability in unitizing acts, has been well documented (Waxler and Mishler, 1966). Errors in identification of acts have been less well studied, although Psathas (1961) has shown that unitizing done while an observer watches group discussion in process can lose as much as 23 percent of the acts produced. Although most of the omissions reported by Psathas seem to have been in the exclusion of nonverbal acts, other results (Riecken, 1958) suggest that omissions may not be uniform across group members. Such inaccuracy in identifying the occurence of units may well obscure whatever regularities exist.

Previously it was mentioned that Stephan and Mishler had reported that an exponential function fit not only the acts data of Bales but even better their own participations data. It has also been mentioned that participations, as Stephan and Mishler defined them, are very close to continuous sequences of acts, what we have called bursts. This similarity between participations and bursts and the fact that the exponential function fits the distribution of participations as well as it does prompts the investigation of bursts in the same way that we have investigated acts, especially since it seems likely that bursts, or changes in actors, would be a much more reliable unit of observation than acts.

Bursts Models

In constructing models for bursts, the simplest possible model is a first order Markov chain, since by definition the speaker must change with each burst. The Horvath process thus requires re-interpretation in order to be applicable to burst data. There are at least three first order models which bear some similarity to the Horvath process. In the first model, H1 , the probability distribution for the next speaker is obtained by eliminating the present speaker and renormalizing the row vector. The other two models, H2 and H3 , eliminate the present speaker and treat the other group members as a new group of size $n-1$ according to the ordinary Horvath model. H2 assumes that it makes no difference who the present speaker is, there being a single, exponentially related distribution of probabilities for the reamining $n-1$ members. H3 , on the other hand, implies that although the distribution of probabilities for the next actor is exponentially related, the nature of the exponential relation may depend on the person who is presently speaking. The first two models involve but one parameter; the third model requires n . Each of these models can be illustrated for four person groups as follows:

$$\begin{bmatrix} o & \frac{ar}{1-a} & \frac{ar^2}{1-a} & \frac{ar^3}{1-a} \\ \frac{a}{1-ar} & o & \frac{ar^2}{1-ar} & \frac{ar^3}{1-ar} \\ \frac{a}{1-ar^2} & \frac{ar}{1-ar^2} & o & \frac{ar^3}{1-ar^2} \\ \frac{a}{1-ar^3} & \frac{ar}{1-ar^3} & \frac{ar^2}{1-ar^3} & o \end{bmatrix} \qquad \begin{bmatrix} o & a & ar & ar^2 \\ a & o & ar & ar^2 \\ a & ar & o & ar^2 \\ a & ar & ar^2 & o \end{bmatrix}$$

H_1 $\qquad$ H_2

$$\begin{bmatrix} o & a_1 & a_1r_1 & a_1r_1^2 \\ a_2 & o & a_2r_2 & a_2r_2^2 \\ a_3 & a_3r_3 & o & a_3r_3^2 \\ a_4 & a_4r_4 & a_4r_4^2 & o \end{bmatrix}$$

H_3

These models have all been called Horvath models since they incorporate an exponential distribution in some manner. If one considers the long run proportion of participations (the fixed probability vector or left eigenvector) it is interesting to observe that the models are not like Horvath's model; they do not have as their equilibrium vector an exponential function. The equilibrium vectors for the three person case for models H1 , H2 , and H3 are:

$$H1 : [(1-a)/c,\ r(1-ar)/c,\ r^2(1-ar^2)/c]$$

$$H2 : [(ar+1)/c,\ (a+r)/c,\ r(a+1)/c]$$

$$H3 : [(1-a_2a_3r_2r_3)/c,\ (1-a_1a_3r_1)/c,\ (1-a_1a_2)/c]$$

where $\underline{c}$ is the appropriate normalizing constant in each case.[4] None of the vectors represent exactly an exponential function, although H1 seems to bear the closest similarity. Nevertheless, because of their use of an exponential function in the transition matrix we shall continue to refer to these three models as Horvath models.

Each of the Horvath models for bursts has a corresponding multinomial generalization. M1 is a renormalization after the elements on the major diagonal have been eliminated, M2 consists of a single set of n-2 parameters which are assumed not to depend on the previous actor, and M3 has n sets of n-2 parameters, one for each possible previous actor. The transition matrices for these models can be illustrated for the four person case as follows:

$$\begin{bmatrix} o & \frac{p_2}{1-p_1} & \frac{p_3}{1-p_1} & \frac{p_4}{1-p_1} \\ \frac{p_1}{1-p_2} & o & \frac{p_3}{1-p_2} & \frac{p_4}{1-p_2} \\ \frac{p_1}{1-p_3} & \frac{p_2}{1-p_3} & o & \frac{p_4}{1-p_3} \\ \frac{p_1}{1-p_4} & \frac{p_2}{1-p_4} & \frac{p_3}{1-p_4} & o \end{bmatrix} \quad \begin{bmatrix} o & p_1 & p_2 & p_3 \\ p_1 & o & p_2 & p_3 \\ p_1 & p_2 & o & p_3 \\ p_1 & p_2 & p_3 & o \end{bmatrix} \quad \begin{bmatrix} o & p_{11} & p_{12} & p_{13} \\ p_{21} & o & p_{22} & p_{23} \\ p_{31} & p_{32} & o & p_{33} \\ p_{41} & p_{42} & p_{43} & o \end{bmatrix}$$

M1 M2 M3

The general case for these models is described in Table 6. Column 2 of Table 6 gives the parameters in general terms. Column 3

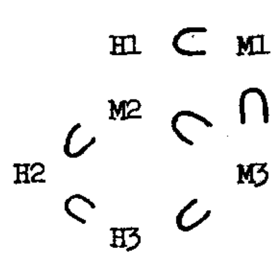
TABLE 6

<u>General Case of the Parameters for the Horvath and Multinomial Burst Models</u>

Models	Parameters	Horvath Case
H1, M1	$m_{ij} = \begin{cases} p_j/(1-p_i) & i \neq j \\ 0 & i = j \end{cases}$	$p_k = ar^{k-1} = \frac{\bar{r}r^{k-1}}{1-r^n}$
H2, M2	$m_{ij} = \begin{cases} p_{j-1} & i < j \\ 0 & i = j \\ p_j & i > j \end{cases}$	$p_k = ar^{k-1} = \frac{\bar{r}r^{k-1}}{1-r^{n-1}}$
H3, M3	$m_{ij} = \begin{cases} p_{ij-1} & i < j \\ 0 & i = j \\ p_{ij} & i > j \end{cases}$	$p_{ik} = a_i r_i^{k-1} = \frac{\bar{r}_i r_i^{k-1}}{1-r_i^{n-1}}$

$\bar{r} = 1-r = \phi$

cites the additional restrictions which are imposed in the case of the Horvath models. The relation of these models is as follows:[5]



The analysis of the model hierarchy follows lines similar to the acts treatment. Maximum **likelihood** estimates of the parameters are used in likelihood ratio tests of goodness of fit for all the inclusions indicated above. The results are shown in Table 7.

TABLE 7

Results of Tests of First Order Horvath and Multinomial Burst Models

Hypothesis	Data Set	LRT	DF	$\simeq$ Sig.
H1 vs. M1	Harvard II-M	225.18	60	$< 10^{-6}$
	Harvard II-F	898.85	108	$< 10^{-6}$
M1 vs. M3	Harvard II-M	268.92	220	.01
	Harvard II-F	613.54	396	$< 10^{-6}$
H2 vs. M2	Harvard II-M	129.00	40	$< 10^{-6}$
	Harvard II-F	652.13	72	$< 10^{-6}$
H2 vs. H3	Harvard II-M	203.22	80	$< 10^{-6}$
	Harvard II-F	318.84	144	$< 10^{-6}$
M2 vs. M3	Harvard II-M	502.76	240	$< 10^{-6}$
	Harvard II-F	950.07	432	$< 10^{-6}$
H3 vs. M3	Harvard II-M	428.54	200	$< 10^{-6}$
	Harvard II-F	1283.37	360	$< 10^{-6}$

The only instance in which the more restrictive model is not clearly rejected is the comparison of M1 and M3 for groups of males, and here the significance level for the aggregated results is .01. This model, M1 , which results from eliminating the major diagonal of the transition matrix for an independent trials multinomial process and renormalizing the row vectors, bears an interesting relation to a model recently put forth by Leik (1965). In that instance Leik was dealing with a probability distribution which resulted from the normalization of certain "tendencies to participate." The two models clearly refer to different units of observation and to different orders of process, but the similarity between the models is interesting.

As before, the aggregation of results from all groups masks some interesting findings. If one treats the groups individually, several of the tests show a rather high proportion of groups for which the more restrictive model is not rejected. Using the .05 level of significance as the criterion for rejection, these results are shown in Table 8. When compared to M3 , it is interesting that M1 ,

TABLE 8

Proportion of Groups Which Do Not Reject the More Restrictive Model at the .05 Level of Significance

Models	DF	Harvard II-M N = 20	Harvard II-F N = 36
H1 vs. M1	3	.60	.25
M1 vs. M3	11	.90	.72
H2 vs. M2	2	.60	.31
H2 vs. H3	4	.65	.64
M2 vs. M3	12	.45	.42
H3 vs. M3	10	.45	.25

which has n-1 parameters, does better than H3 , which has n parameters. Obviously, it is not only the number of parameters, but the manner in which they are incorporated which determines how restrictive a model is.

None of the suggested Horvath models did very well against its corresponding multinomial model.

To provide a fuller picture of the results of testing M1 against M3 , a complete enumeration of the results for the Harvard II-M data set is shown in Table 9. The tests are presented so that

TABLE 9

Results of Testing M1 vs M3, Group by Group
Harvard II-M Data, DF = 11

Group	Session	LRT	≃ Sig.
1	1	19.41	.05
	3	9.83	.55
2	1	7.14	.79
	3	14.13	.23
3	1	13.67	.25
	3	15.28	.17
4	1	9.38	.59
	3	5.26	.92
5	1	11.97	.37
	3	21.41	.03
6	1	11.06	.44
	3	15.69	.15
7	1	10.17	.52
	3	18.53	.07
8	1	6.22	.86
	3	19.42	.05
9	1	8.95	.63
	3	24.94	.009
10	1	7.98	.72
	3	18.48	.07

one can see the results for both the first session and the third session (the second task oriented session) for each group. In only one case is a result significant at less than the .01 level. Finally, one will notice that in eight of the ten groups the value of the test statistic is larger in the third session than it is in the first session. If the probability of a larger value in the third session were 1/2 and the outcomes for different groups were independent, then the probability of eight or more events of this type would be .055. Thus, there appears to be a slight tendency for the process to become less like model M1 as the groups proceed from session one to session three.

Finally, we wish to consider whether a higher order chain would do much better in describing the process. Again because of the rapid increase in the number of parameters for higher order processes we shall have to limit our analysis to a comparison of first and second order chains. The model hierarchy which we shall test is as follows:

$$\begin{array}{ccccc} M3 & = & M3_{D,1} & \subset & M3_{D,2} \\ & & \cup & & \\ & & M3_{S,1} & \subset & M3_{S,2} \end{array}$$

The least restrictive first order model, $M3_{D,1}$, assumes that each group can be characterized by some transition matrix with a zero major diagonal. This is the same as the previous model M3 . $M3_{S,1}$ assumes that every group can be characterized by the same transition matrix; $M3_{D,2}$ and $M3_{S,2}$ are the second order equivalents.

The results are contained in Table 10. All aggregate tests reject the more restrictive and the lower order models. In the analysis of $M3_{D,1}$ vs. $M3_{D,2}$ group by group, in only 3 of the 20 sessions of males and 13 of the 36 sessions of females was the first order model

TABLE 10

Results of Tests of First Order versus Second Order Multinomial Burst Models

Hypothesis	Data Set	LRT	DF	≃ Sig.
$M3_{S,1}$ vs $M3_{D,1}$	Harvard II-M	664.80	285	$< 10^{-6}$
	Harvard II-F	2019.52	525	$< 10^{-6}$
$M3_{S,1}$ vs $M3_{S,2}$	Harvard II-M	651.87	45	$< 10^{-6}$
	Harvard II-F	830.23	45	$< 10^{-6}$
$M3_{S,2}$ vs $M3_{D,2}$	Harvard II-M	1626.88	1140	$< 10^{-6}$
	Harvard II-F	3726.06	2100	$< 10^{-6}$
$M3_{D,1}$ vs $M3_{D,2}$	Harvard II-M	1615.14	900	$< 10^{-6}$
	Harvard II-F	2541.97	1620	$< 10^{-6}$

not rejected at the .05 level. At the .001 level, however, the lower order model was not rejected in 12 (60%) of the sessions of males and in 24 (67%) of the sessions of females.

Discussion

Throughout the tests of both the acts models and the bursts models, the results have been associated with levels of significance at or beyond what is typically adopted as the criterion for rejection. In Table 1, for example, the results of the tests of H_S vs. M_S were significant at the .001, .20, .0002, .003, and 10^{-6} levels, all but the second of which would ordinarily be associated with rejection of the hypothesis, not with acceptance. That particular test compared the aggregated, ordered data to an exponential function and is probably the test that would have been done had Stephan and Mishler subjected their distributions to a statistical test. Recalling that it was the fit (in percentage error) of the exponential to the distribution of aggregated, ordered units which motivated interest in this area of research, it is perhaps surprising that the tests appear to indicate rejection.

In the beginning of this paper we mentioned Coleman's comment that with a large amount of data, rejection can occur even if a model fits data very well. Regardless of how slight the difference between the true value of a quantity and the hypothesized value of that quantity, with sufficient data gathered in the appropriate manner, any good statistical test will lead to rejection of the hypothesized value.

In the present case the amount of data ranges from 11,410 acts in the Yale data to 39,945 in the Chicago data. The aggregated data from a given set differ from estimates based on the exponential function by .5, .3, .6, .4, and 1.6 percent for the Chicago, Yale, Harvard I, Harvard II-M, and Harvard II-F data sets respectively (Kadane and Lewis, 1968). For p_1 and p_2 , the closeness of the aggregated data to the estimates can be seen in Graphs 1 - 5. Despite the fact that the aggregated data appear to be close to the Horvath predictions, the tests, H_S vs. M_S , indicate the extremely small chance that the data would be even this far away if the Horvathian model were correct. On the other hand, many of the points, $(\hat{p}_1, \hat{p}_2)$, which are further away from the Horvath line are involved in tests, H_D vs. M_D , which do not show rejection. The reason for this "paradox" is that, compared to M_D , the single point estimated by M_S has much more data. Therefore, its variance is much less, and the region of non-rejection at any level of significance is much smaller than for the M_D points.

It is because of this relation between the amount of data and the power of the test that we initially indicated that we viewed the present statistical testing as but one way of evaluating the goodness of these models. Certainly we do not always expect that our hypotheses are true -- especially when they are "straw men," and the hypothesis that there is no difference whatsoever between two populations is usually a "straw man"; such expectations would make no more sense in the social sciences than in physical sciences. Alternatively, it would seem reasonable to develop methods of analyzing models which will allow one to make estimates of the magnitude of error he will incur in using the model in various ways. At the present time an alternative of this type is being investigated, but as a first step in the evaluation of the present stochastic models, we have chosen to present the results of the standard approach to the evaluation of goodness of fit.

<u>Summary</u>

When Horvath published his theory of participation in group discussions, the only consequence with which he concerned himself was the prediction of the total distribution of participation, and in this respect the model and the data appeared in accord. Because the model appeared to fit in this one area it seemed reasonable to attempt a more thorough evaluation of the model.

The first evidence consisted of graphs of the space for parameters p_1 and p_2. The space for the Horvath models is a curved line through the space for the multinomial models. The data fell on both sides of the line and in a widely dispersed area. When the data are aggregated, however, the aggregate values fell fairly close to the Horvath line.

In the statistical tests of the acts models, the Horvath model with a separate transition matrix for each group did fairly well against the multinomial model with a separate transition matrix for each group, H_D vs. M_D, although this was true only for the groups of males. At the levels of significance investigated, the test failed to show rejection for twice the proportion of groups of males as it did for groups of females.

When the data from the groups were aggregated and a single parameter set for the Horvath model was tested against a single parameter set for the multinomial model, H_S vs. M_S, the significance levels ranged from .20 to 10^{-6}, again with the more restrictive model fitting the data from males better than the data from females. Although the results were mixed, all but one of the results were at a level of significance which would ordinarily be associated with rejection of the hypothesis. All other zero order tests showed strong rejection.

When zero and first order processes were compared to first and second order processes, the higher orders did considerably better, although for about 60 percent of the groups, when the first and second orders were compared, the first order was not rejected at the .001 level. Consequently, any conclusions about the goodness of fit of lower order models, e.g. H_D vs. M_D, must be made with reservations; the results of such a comparison may indicate only that there is not much difference in the two models, both of which may have a poor fit to the data.

In the statistical results of the bursts models, none of the Horvathian models did very well against its multinomial generalization. The multinomial model which resulted from eliminating the major diagonal of a zero order model and then renormalizing the row vectors, M1, performed fairly well in comparison to the more general model, M3. Model M1 performed better than model H3 which has one more parameter.

The results of testing the order of the process for the bursts models were comparable to those for the acts models: using the .001 level of significance, the first order process was not rejected when compared to the second order process for 60 percent of the groups.

Finally, large amounts of data were used in testing the models. Some of the more restrictive models which were rejected on the basis of the outcomes of the statistical tests may be associated with relatively small errors in making predictions about particular quantities of interest. For this reason, we would suggest that the more restrictive models should not necessarily be rejected at the present time solely on the basis of the outcomes of the statistical tests.

FOOTNOTES

[1]The likelihood function and a demonstration that there exists a unique real root, which lies in the interval [0,1] , appear in the Appendix.

[2]These data are from unpublished work done under Family Interaction Studies Using Revealed Difference, NIMH Grant No. MH05572-03, in the Social Psychology Laboratory, at the University of Chicago.

[3]The construction of these and subsequently mentioned tests is given in the Appendix.

[4]The derivation of these vectors is given in the Appendix.

[5]If $n = 3$, then $H_2 = M_2$, and $H_3 = M_3$.

REFERENCES

Anderson, T.W. and L.A. Goodman (1957). "Statistical Inference about Markov Chains." Annals of Mathematical Statistics, 28: 89-110.

Bales, R. (1951). Interaction Process Analysis. Reading, Massachusetts: Addison-Wesley.

Bales, R., F. Strodtbeck, T. Mills, and M. Roseborough (1951). "Channels of Communication in Small Groups." American Sociological Review, 16: 461-468.

Bonacich, P. (1968). "Specialization and Differentiation in Small Laboratory Groups." Ph.D. Dissertation, Harvard University.

Coleman, J. (1960). "The Mathematical Study of Small Groups." pp. 1-149 in Herbert Solomon (ed.), Mathematical Thinking in the Measurement of Behavior. Glencoe, Illinois: The Free Press of Glencoe, Illinois.

Horvath, W. (1965). "A Mathematical Model of Participation in Small Group Discussions." Behavioral Science, 10: 164-166.

Kadane, J. and G. Lewis (1968). "Alternative Explanations of Regularity in the Distribution of Participation." In preparation.

Kunz, K. (1957). Numerical Analysis. New York: McGraw-Hill.

Leik, R. (1965). "Type of Group and the Probability of Initiating Acts." Sociometry, 28: 57-65.

________ (1967). "The Distribution of Acts in Small Groups." Sociometry, 30: 280-299.

Olmsted, M. (1954). "Orientation and Role in the Small Group." American Sociological Review, 19: 741-751.

Psathas, G. (1961). "Alternative Methods for Scoring Interaction Process Analysis." Journal of Social Psychology, 53: 97-103.

Riecken, H. (1958). "The Effect of Talkativeness on Ability to Influence Group Solutions of Problems." Sociometry, 21: 309-321.

Stephan, F. (1952). "The Relative Rate of Communication between Members of Small Groups." American Sociological Review, 17: 482-486.

Stephan, F. and E. Mishler (1952). "The Distribution of Participation in Small Groups: An Exponential Approximation." _American Sociological Review_, 17: 598-608.

Waxler, N. and E. Mishler (1966). "Scoring and Reliability Problems in Interaction Process Analysis: A Methodological Note." _Sociometry_, 29: 28-40.

APPENDIX

TO

STOCHASTIC THEORIES OF PARTICIPATION

TABLE OF CONTENTS

APPENDIX A

THE HORVATH LIKELIHOOD FUNCTION

The likelihood function for the Horvath model is given by

$$l(\varphi) = \prod_{i=1}^{n} [ar^{i-1}]^{N_i} = a^{\Sigma N_i} r^{\Sigma(i-1)N_i}$$

$$= a^{N.} r^{\Sigma iN_i - N.} = a^{N.} r^{M-N.}$$

where $\{N_i\}$ are the ordered, aggregated frequency data. Since $r < 1$, $l(\varphi)$ is clearly maximized by choosing the ordering of the N_i which minimizes $M = \Sigma iN_i$. This ordering is exactly the ranking from high to low. Substituting for a and r, we have

$$l(\varphi) = \left[\frac{1 - \bar{\varphi}}{1 - \bar{\varphi}^n}\right]^{N.} \bar{\varphi}^{M-N.} ,$$

where $\bar{\varphi} = 1 - \varphi$, $\varphi \neq 0, 1$. Maximize the function

$$g(\bar{\varphi}) = \left[\frac{1 - \bar{\varphi}}{1 - \bar{\varphi}^m}\right]^{\alpha} \bar{\varphi}^{\beta} .$$

$$h(\bar{\varphi}) = \ln g(\bar{\varphi}) = \alpha \ln(1 - \bar{\varphi}) - \alpha \ln(1 - \bar{\varphi}^m) + \beta \ln \bar{\varphi}$$

$$h'(\bar{\varphi}) = -\alpha/(1 - \bar{\varphi}) + \alpha/(1 - \bar{\varphi}^m) m\bar{\varphi}^{m-1} + \beta/\bar{\varphi}$$

$$= \frac{-\alpha(1 - \bar{\varphi}^m)\bar{\varphi} + \alpha(1 - \bar{\varphi})\bar{\varphi} m\bar{\varphi}^{m-1} + \beta(1 - \bar{\varphi})(1 - \bar{\varphi}^m)}{(1 - \bar{\varphi})(1 - \bar{\varphi}^m)\bar{\varphi}} = 0$$

$$\bar{\varphi}^{m}[\alpha m(1-\bar{\varphi}) + \alpha\bar{\varphi} - \beta(1-\bar{\varphi})] + [-\alpha\bar{\varphi} + \beta(1-\bar{\varphi})] = 0$$

$$= \bar{\varphi}^{m}[\bar{\varphi}(\beta - (m-1)\alpha) + (m\alpha-\beta)] - [\bar{\varphi}(\alpha+\beta) - \beta]$$

$$= (\bar{\varphi}-1)[(\beta - (m-1)\alpha)\bar{\varphi}^{m} + \alpha\bar{\varphi}^{m-1} + \dots + \alpha\bar{\varphi} - \beta]$$

$$= (\bar{\varphi}-1)^{2}[(\beta - (m-1)\alpha)\bar{\varphi}^{m-1} + (\beta - (m-2)\alpha)\bar{\varphi}^{m-2} + \dots + (\beta-\alpha)\bar{\varphi}+\beta]$$

Now the problem has been reduced to finding the roots of a polynomial of degree $(m-1)$. For the special case $l(\varphi)$,

$$\alpha = N.$$

$$\beta = M-N.$$

$$m = n$$

and the polynomial becomes

$$f(\bar{\varphi}) = \bar{\varphi}^{n-1}[M - nN.] + \bar{\varphi}^{n-2}[M - (n-1)N.] + \dots + \bar{\varphi}[M - 2N.] + [M - N.]$$

Now note that

$$N. \leq M \leq nN.$$

The first equality holds **iff** $N_1 = N.$; the second **iff** $N_n = N.$. The maximum likelihood ordering of the N_i gives a better upper bound. When the N_i are ranked from highest to lowest, the maximum possible M occurs when

$$N_1 = N_2 = \dots = N_n = c$$

Then $M = \Sigma iN_i = cn(n+1)/2 = N.(n+1)/2$, since $N. = cn$. The resulting inequality is

$$N. \leq M \leq (n+1)/2 \cdot N.$$

But this implies that the coefficients of f are all negative up to some term, and then all the remaining coefficients are positive. Therefore, by Descartes rule of signs,[1] the number of positive roots must be ≤ 1, and must be smaller by a multiple of two. Thus there exists a unique positive root.

As $\bar{\varphi} \to \infty$, the first term in $f(\bar{\varphi})$ eventually dominates, and since this coefficient is negative,

$$f(\infty) < 0 .$$

But

$$f(1) = n[M - N.] - n(n-1)/2 \cdot [N.] = n[M - (n+1)/2 \cdot N.] \leq 0$$

by the inequality. Thus there are no roots between 1 and ∞ and we have $0 \leq \bar{\varphi} \leq 1$.

The degenerate cases $\bar{\varphi} = 0, 1$ have special meaning. When $\bar{\varphi} = 0$, the highest participant is the only participant; when $\bar{\varphi} = 1$, the participation probabilities are identical for every group member.

---

[1]See K.S. Kunz, Numerical Analysis, New York: McGraw-Hill, 1957, p. 18.

APPENDIX B

TESTS OF THE MARKOV ORDER 0 ACT MODELS

Hypothesis	Parameters	Dimension	DF	Notes
H_S vs H_D	φ $\{\varphi_k\}$	1 s	s-1	$[N_i(k)]$: Markov 0 frequency matrix for group k . $i = 1, \ldots, n$ member (numbered from highest participant to lowest) $k = 1, \ldots, s$ group $[h_i]$: parameters of H_S $h_i(k)$: parameters of H_D (for estimation see Appendix A).

$$\lambda = \prod_{i,k} [\hat{h}_i/\hat{h}_i(k)]^{N_i(k)}$$

$$\chi^2 = \sum_{i,k} N_i(k)[\hat{h}_i(k) - \hat{h}_i]^2/\hat{h}_i$$

H_S vs M_S	φ $\{m_i\}$	1 (n-1)	n-2	$\hat{m}_i = N_i(\cdot)/N.(\cdot)$

$$\lambda = \prod_i [\hat{h}_i/\hat{m}_i]^{N_i(\cdot)}$$

$$\chi^2 = \sum_i N.(\cdot)[\hat{m}_i - \hat{h}_i]^2/\hat{h}_i$$

H_S vs M_D	φ $\{m_i(k)\}$	1 (n-1)s	(n-1)s-1	

$$\lambda = \prod_{i,k} [\hat{h}_i/\hat{m}_i(k)]^{N_i(k)}$$

$$\chi^2 = \sum_{i,k} N.(k)[\hat{m}_i(k) - \hat{h}_i]^2/\hat{h}_i$$

$m_i(k) = N_i(k)/N.(k)$

H_D vs M_D	$\{\varphi_k\}$ $\{m_i(k)\}$	s (n-1)s	(n-2)s	

$$\lambda = \prod_{i,k} [\hat{h}_i(k)/\hat{m}_i(k)]^{N_i(k)}$$

$$\chi^2 = \sum_{i,k} N.(k)[\hat{m}_i(k) - \hat{h}_i(k)]^2/\hat{h}_i(k)$$

Hypothesis	Parameters	Dimension	DF	Notes
M_S vs M_D	$\{m_i\}$	m-1	(n-1)(s-1)	
	$\{m_i(k)\}$	(n-1)s		

$$\lambda = \prod_{i,k} [\hat{m}_i/\hat{m}_i(k)]^{N_i(k)}$$

$$\chi^2 = \sum_{i,k} N.(k)[\hat{m}_i(k) - \hat{m}_i]^2/\hat{m}_i$$

APPENDIX C

TESTS OF THE MARKOV ORDER 0, 1, AND 2 ACT MODELS

Hypothesis	Parameters	Dimension	DF	Notes
$M_{S,0}$ vs $M_{S,1}$	$\{p_j\}$ $\{p_{ij}\}$	$n-1$ $n(n-1)$	$(n-1)^2$	$[N_{ij}(g)]$: Markov 1 frequency matrix for group g . $i,j = 1, \ldots, n$ member (numbered from highest participant to lowest) $g = 1, \ldots, s$ group

$$\lambda = \prod_{i,j} [\hat{p}_j/\hat{p}_{ij}]^{N_{ij}(\cdot)}$$

$\hat{p}_j = N_{\cdot j}(\cdot)/N_{\cdot\cdot}(\cdot)$

$$\chi^2 = \sum_{i,j} N_{i\cdot}(\cdot)[\hat{p}_{ij} - \hat{p}_j]^2/\hat{p}_j$$

$\hat{p}_{ij} = N_{ij}(\cdot)/N_{i\cdot}(\cdot)$

Hypothesis	Parameters	Dimension	DF
$M_{S,0}$ vs $M_{D,0}$	$\{p_i\}$ $\{p_i(g)\}$	$n-1$ $(n-1)s$	$(n-1)(s-1)$

$$\lambda = \prod_{g,i} [\hat{p}_i/\hat{p}_i(g)]^{N_i(g)}$$

$\hat{p}_i = N_i(\cdot)/N_\cdot(\cdot)$

$$\chi^2 = \sum_{g,i} N_\cdot(g)[\hat{p}_i(g) - \hat{p}_i]^2/\hat{p}_i$$

$\hat{p}_i(g) = N_i(g)/N_\cdot(g)$

Hypothesis	Parameters	Dimension	DF
$M_{D,0}$ vs $M_{D,1}$	$\{p_j(g)\}$ $\{p_{ij}(g)\}$	$(n-1)s$ $n(n-1)s$	$(n-1)^2 s$

$$\lambda = \prod_{g,i,j} [\hat{p}_j(g)/\hat{p}_{ij}(g)]^{N_{ij}(g)}$$

$\hat{p}_j(g) = N_{\cdot j}(g)/N_{\cdot\cdot}(g)$

$$\chi^2 = \sum_{g,i,j} N_{i\cdot}(g)[\hat{p}_{ij}(g) - \hat{p}_j(g)]^2/\hat{p}_j(g)$$

$\hat{p}_{ij}(g) = N_{ij}(g)/N_{i\cdot}(g)$

Hypothesis	Parameters	Dimension	DF	Notes
$M_{S,1}$ vs $M_{D,1}$	$\{p_{ij}\}$	$n(n-1)$	$n(n-1)(s-1)$	
	$\{p_{ij}(g)\}$	$n(n-1)s$		

$$\lambda = \prod_{g,i,j} [\hat{p}_{ij}/\hat{p}_{ij}(g)]^{N_{ij}(g)}$$

$$\chi^2 = \sum_{g,i,j} N_{i\cdot}(g)[\hat{p}_{ij}(g) - \hat{p}_{ij}]^2/\hat{p}_{ij}$$

Hypothesis	Parameters	Dimension	DF	Notes
$M_{S,1}$ vs $M_{S,2}$	$\{p_{jk}\}$	$n(n-1)$	$n(n-1)^2$	$[N_{ijk}(g)]$: Markov 2 frequency matrix for group g .
	$\{p_{ijk}\}$	$n^2(n-1)$		

$$\lambda = \prod_{i,j,k} [\hat{p}_{jk}/\hat{p}_{ijk}]^{N_{ijk}(\cdot)}$$

$\hat{p}_{jk} = N_{\cdot jk}(\cdot)/N_{\cdot j\cdot}(\cdot)$

$$\chi^2 = \sum_{i,j,k} N_{ij\cdot}(\cdot)[\hat{p}_{ijk}-\hat{p}_{jk}]^2/\hat{p}_{jk}$$

$\hat{p}_{ijk} = N_{ijk}(\cdot)/N_{ij\cdot}(\cdot)$

Hypothesis	Parameters	Dimension	DF	Notes
$M_{D,1}$ vs $M_{D,2}$	$\{p_{jk}(g)\}$	$n(n-1)s$	$n(n-1)^2 s$	
	$\{p_{ijk}(g)\}$	$n^2(n-1)s$		

$$\lambda = \prod_{g,i,j,k} [\hat{p}_{jk}(g)/\hat{p}_{ijk}(g)]^{N_{ijk}(g)}$$

$\hat{p}_{jk}(g) = N_{\cdot jk}(g)/N_{\cdot j\cdot}(g)$

$$\chi^2 = \sum_{g,i,j,k} N_{ij\cdot}(g)[\hat{p}_{ijk}(g)-\hat{p}_{jk}(g)]^2/\hat{p}_{jk}(g)$$

$\hat{p}_{ijk}(g) = N_{ijk}(g)/N_{ij\cdot}(g)$

Hypothesis	Parameters	Dimension	DF	Notes
$M_{S,2}$ vs $M_{D,2}$	$\{p_{ijk}\}$	$n^2(n-1)$	$n^2(n-1)(s-1)$	
	$\{p_{ijk}(g)\}$	$n^2(n-1)s$		

$$\lambda = \prod_{g,i,j,k} [\hat{p}_{ijk}/\hat{p}_{ijk}(g)]^{N_{ijk}(g)}$$

$$\chi^2 = \sum_{g,i,j,k} N_{ij\cdot}(g)[\hat{p}_{ijk}(g) - \hat{p}_{ijk}]^2/\hat{p}_{ijk}$$

APPENDIX D

LIKELIHOOD FUNCTIONS FOR THE MARKOV ORDER 1 BURST MODELS

$$M_1: \ \ell = \prod_{i \neq j} \left[\frac{p_j}{1 - p_i}\right]^{N_{ij}} = \frac{\prod_{i \neq j} p_j^{N_{ij}}}{\prod_{i \neq j} (1 - p_i)^{N_{ij}}} = \frac{\prod_j p_j^{N_{\cdot j}}}{\prod_i (1 - p_i)^{N_{i\cdot}}}$$

where $[N_{ij}]$ is the Markov 1 burst frequency matrix.

$$L = \ln \ell = \sum_k [N_{\cdot k} \ln p_k - N_{k\cdot} \ln(1 - p_k)] + \lambda[1 - \sum_k p_k]$$

$$\partial L/\partial p_k = N_{\cdot k}/p_k + N_{k\cdot}/(1 - p_k) - \lambda = 0$$

$$N_{\cdot k}(1 - p_k) + N_{k\cdot} p_k = \lambda p_k(1 - p_k)$$

$$p_k^2\lambda + p_k(N_{k\cdot} - N_{\cdot k} - \lambda) + N_{\cdot k} = 0$$

$$p_k = [(\lambda + N_{\cdot k} - N_{k\cdot}) \pm \{(\lambda + N_{\cdot k} - N_{k\cdot})^2 - 4\lambda N_{\cdot k}\}^{1/2}]/(2\lambda)$$

$$= [(\lambda + N_{\cdot k} - N_{k\cdot}) \pm \{(\lambda - N_{\cdot k} - N_{k\cdot})^2 - 4N_{\cdot k}N_{k\cdot}\}^{1/2}]/(2\lambda)$$

Constraints:

$$0 \leq p_k \leq 1$$

$$1 = \sum_1^n p_k = 1/(2\lambda)[n\lambda + \sum_k \pm \{(\lambda - N_{\cdot k} - N_{k\cdot})^2 - 4N_{\cdot k}N_{k\cdot}\}^{1/2}]$$

$$= n/2 + \frac{1}{(2\lambda)} \sum_k \pm \{(\lambda - N_{\cdot k} - N_{k\cdot})^2 - 4N_{\cdot k}N_{k\cdot}\}^{1/2}$$

$$(n-2)\lambda = \sum_{k} \mp \{(\lambda - N_{\cdot k} - N_{k\cdot})^2 - 4N_{\cdot k}N_{k\cdot}\}^{1/2}$$

$$0 \leq p_n \leq p_{n-1} \leq \cdots \leq p_1 \leq 1$$

since the square root must be real,

$$(\lambda - N_{\cdot k} - N_{k\cdot})^2 \geq 4N_{\cdot k}N_{k\cdot}$$

$$|\lambda - N_{\cdot k} - N_{k\cdot}| \geq 2(N_{\cdot k}N_{k\cdot})^{1/2}$$

if $\lambda \geq N_{\cdot k} + N_{k\cdot}$

$$\lambda - N_{\cdot k} - N_{k\cdot} \geq 2(N_{\cdot k}N_{k\cdot})^{1/2}$$

$$\lambda \geq N_{\cdot k} + N_{k\cdot} + 2(N_{\cdot k}N_{k\cdot})^{1/2} = (N_{\cdot k}^{1/2} + N_{k\cdot}^{1/2})^2$$

if $\lambda \leq N_{\cdot k} + N_{k\cdot}$

$$N_{\cdot k} + N_{k\cdot} - \lambda \geq 2(N_{\cdot k}N_{k\cdot})^{1/2}$$

$$\lambda \leq N_{\cdot k} + N_{k\cdot} - 2(N_{\cdot k}N_{k\cdot})^{1/2} = (N_{\cdot k}^{1/2} - N_{k\cdot}^{1/2})^2$$

therefore either

$$\lambda \leq (N_{\cdot k}^{1/2} - N_{k\cdot}^{1/2})^2$$

or

$$\lambda \geq (N_{\cdot k}^{1/2} + N_{k\cdot}^{1/2})^2$$

Also, $N_{\cdot k} = N_{k\cdot}$ for at least one value of k for $n \geq 3$. Since the λ is the same for all k,

$$\lambda \leq \min_k (N_{\cdot k}^{1/2} - N_{k\cdot}^{1/2})^2 = 0$$

or

$$\lambda \geq \max_k (N_{\cdot k}^{1/2} + N_{k\cdot}^{1/2})^2 = (N_{\cdot 1}^{1/2} + N_{1\cdot}^{1/2})^2$$

Now $N_{\cdot k} \neq N_{k\cdot}$ iff person k began or ended the acts sequence (not both). In this case $N_{\cdot k} = N_{k\cdot} \pm 1$.

Consider the case

$$p_k = [(\lambda + N_{\cdot k} - N_{k\cdot}) + ((\lambda + N_{\cdot k} - N_{k\cdot})^2 - 4\lambda N_{\cdot k})^{1/2}]/(2\lambda)$$

If $N_{\cdot k} = N_{k\cdot}$,

$$p_k = 1/2 + (\lambda^2 - 4\lambda N_{\cdot k})^{1/2}/(2\lambda) = 1/2 + (\lambda - 4N_{\cdot k})^{1/2}/(2\lambda^{1/2}) > 1/2$$

for $\lambda \geq (N_{\cdot 1}^{1/2} + N_{1\cdot}^{1/2})^2$.

If $N_{\cdot k} = N_{k\cdot} \pm 1$,

$$p_k = 1/2 \pm 1/(2\lambda) + 1/(2\lambda)((\lambda \pm 1)^2 - 4\lambda N_{\cdot k})^{1/2}$$

$$= 1/2 + 1/(2\lambda)[(\lambda^2 \pm 2\lambda + 1 - 4\lambda N_{\cdot k})^{1/2} \pm 1]$$

$$p_k > 1/2 + 1/(2\lambda)[\lambda^{1/2}(\lambda - 4N_{\cdot k} \pm 2)^{1/2} \pm 1] > 1/2$$

for $\lambda > (N_{\cdot 1}^{1/2} + N_{1\cdot}^{1/2})^2$, $k \neq 1$.

Since $n \geq 3$, there can be no more than one (+) chosen among the $\{p_i\}$. If there is one at all, it must be chosen for p_1.

$$H_1: \quad l(\varphi) = \frac{\prod_k [ar^{k-1}]^{N_{\cdot k}}}{\prod_k [1-ar^{k-1}]^{N_{k\cdot}}} = \frac{a^{N_{\cdot\cdot}} r^{\Sigma kN_{\cdot k}-N_{\cdot\cdot}}}{\prod_k [1-ar^{k-1}]^{N_{k\cdot}}} = \frac{\left[\frac{1-\bar\varphi}{1-\bar\varphi^n}\right]^{N_{\cdot\cdot}} \bar\varphi^{\Sigma kN_{\cdot k}-N_{\cdot\cdot}}}{\prod_k \left[1 - \frac{1-\bar\varphi}{1-\bar\varphi^n}\cdot\bar\varphi^{k-1}\right]^{N_{k\cdot}}}$$

$$= \frac{\left(\frac{1-\bar\varphi}{1-\bar\varphi^n}\right)^{N_{\cdot\cdot}} \bar\varphi^{\Sigma kN_{\cdot k}-N_{\cdot\cdot}}}{\prod_k \left[\frac{(1-\bar\varphi^n) - (1-\bar\varphi)\bar\varphi^{k-1}}{1-\bar\varphi^n}\right]^{N_{k\cdot}}} = \frac{(1-\bar\varphi)^{N_{\cdot\cdot}} \bar\varphi^{\Sigma kN_{\cdot k}-N_{\cdot\cdot}}}{\prod_k [(1-\bar\varphi^n) - (1-\varphi)\bar\varphi^{k-1}]^{N_{k\cdot}}}$$

$$= \frac{\bar\varphi^{\Sigma kN_{\cdot k}-N_{\cdot\cdot}}}{\prod_k [\sum_1^n \bar\varphi^{i-1} - \bar\varphi^{k-1}]^{N_{k\cdot}}}$$

$$L(\bar\varphi) = (\Sigma kN_{\cdot k} - N_{\cdot\cdot})\ln\bar\varphi - \sum_k N_{k\cdot} \ln\Big(\sum_{i\neq k} \bar\varphi^{i-1}\Big)$$

$$\frac{\partial L}{\partial\bar\varphi} = \frac{\Sigma kN_{\cdot k} - N_{\cdot\cdot}}{\bar\varphi} - \sum_k N_{k\cdot} \frac{\sum_{i\neq k}(i-1)\bar\varphi^{i-2}}{\sum_{i\neq k}\bar\varphi^{i-1}} = 0$$

$$= \Sigma kN_{\cdot k} - N_{\cdot\cdot} - \sum_k N_{k\cdot}\Big(\sum_{i\neq k} i\bar\varphi^{i-1} - \sum_{i\neq k}\bar\varphi^{i-1}\Big)\Big/\Big(\sum_{i\neq k}\bar\varphi^{i-1}\Big)$$

$$= \Sigma kN_{\cdot k} - \sum_k N_{k\cdot}\left(\frac{\sum_{i\neq k} i\bar\varphi^{i-1}}{\sum_{i\neq k}\bar\varphi^{i-1}}\right) = \sum_k \left[\frac{kN_{\cdot k}\sum_{i\neq k}\bar\varphi^{i-1} - N_{k\cdot}\sum_{i\neq k} i\bar\varphi^{i-1}}{\Sigma\bar\varphi^{i-1}}\right]$$

$$= \sum_k \left[\frac{\sum_{i\neq k}(kN_{\cdot k} - iN_{k\cdot})\bar\varphi^{i-1}}{\sum_{i\neq k}\bar\varphi^{i-1}}\right] = 0$$

$$M_2: \quad 1 = \prod_{j=1}^{n-1} q_j^{\sum_{i\leq j} N_{ij+1} + \sum_{i>j} N_{ij}} = \prod_{j=1}^{n-1} q_j^{\alpha_j}$$

$$\hat{q}_j = \alpha_j / N_{\cdot\cdot}$$

$$H_2: \quad l(\bar{\varphi}) = \prod_{j=1}^{n-1} [ar^{j-1}]^{\sum_{i\leq j} N_{ij+1} + \sum_{i>j} N_{ij}} = a^{N_{\cdot\cdot}} r^{\sum_j j\alpha_j - N_{\cdot\cdot}} = a^{N_{\cdot\cdot}} r^{\sum_j N_{\cdot j} - N_{\cdot\cdot} - \sum_i \sum_{j>i} N_{ij}}$$

The maximum likelihood estimates are obtained by setting

$$\alpha = N_{\cdot\cdot}$$

$$\beta = \sum_j N_{\cdot j} - N_{\cdot\cdot} - \sum_i \sum_{j>i} N_{ij}$$

$$m = n-1$$

in the function $g(\bar{\varphi})$ (Appendix A).

$$M_3: \quad 1 = \prod_{i\neq j} p_{ij}^{N_{ij}} \qquad \hat{p}_{ij} = N_{ij}/N_{i\cdot}$$

$$H_3: \quad l(\bar{\varphi}_i) = \prod_{j<i} [a_i r_i^{j-1}]^{N_{ij}} \prod_{j\geq i} [a_i r_i^{j-1}]^{N_{ij+1}}$$

$$= a_i^{\sum_{j<i} N_{ij} + \sum_{j\geq i} N_{ij+1}} \, r_i^{\sum_{j<i} (j-1)N_{ij} + \sum_{j\geq i} (j-1)N_{ij+1}}$$

$$= a_i^{N_{i\cdot}} \, r_i^{\sum_{j<i} jN_{ij} + \sum_{j\geq i} jN_{ij+1} - N_{i\cdot}}$$

$$= a_i^{N_{i\cdot}} \, r_i^{\sum_{j=1}^{n} jN_{ij} - \sum_{j>i} N_{ij} - N_{i\cdot}}$$

The maximum likelihood estimates are obtained by setting

$$\alpha = N_i.$$

$$\beta = \sum_j jN_{ij} - \sum_{j>i} N_{ij} - N_i.$$

$$m = n-1$$

in $g(\overline{\varphi})$ (Appendix A).

APPENDIX E

TESTS OF THE MARKOV ORDER 1 BURST MODELS

Hypothesis	Parameters	Dimension	DF	Notes
H1 vs M1	φ	1	n-2	$[N_{ij}]$: Markov 1 burst frequency matrix i = 1, ..., n member (numbered from highest participant to lowest).
	$\{m_{ij}(1)\}$	n-1		$\{h_{ij}(K)\}$: parameters of H_K (for estimation see Appendix D).

$$\lambda = \prod_{i,j} [\hat{h}_{ij}(1)/\hat{m}_{ij}(1)]^{N_{ij}}$$

$$\chi^2 = \sum_{i,j} N_i \cdot [\hat{m}_{ij}(1) - \hat{h}_{ij}(1)]^2/\hat{h}_{ij}(1)$$

Hypothesis	Parameters	Dimension	DF
M1 vs M3	$\{m_{ij}(1)\}$	n-1	n(n-2)-(n-1)
	$\{m_{ij}(3)\}$	n(n-2)	

$$\lambda = \prod_{i,j} [\hat{m}_{ij}(1)/\hat{m}_{ij}(3)]^{N_{ij}}$$

$$\chi^2 = \sum_{i,j} N_i \cdot [\hat{m}_{ij}(3) - \hat{m}_{ij}(1)]^2/\hat{m}_{ij}(1)$$

Hypothesis	Parameters	Dimension	DF
H2 vs M2	φ	1	n-3
	$\{m_{ij}(2)\}$	n-2	

$$\lambda = \prod_{i,j} [\hat{h}_{ij}(2)/\hat{m}_{ij}(2)]^{N_{ij}}$$

$$\chi^2 = \sum_{i,j} N_i \cdot [\hat{m}_{ij}(2) - \hat{h}_{ij}(2)]^2/\hat{h}_{ij}(2)$$

Hypotheses	Parameters	Dimension	DF	Notes
H2 vs H3	φ	1	n-1	
	$\{\varphi_i\}$	n		

$$\lambda = \prod_{i,j} [\hat{h}_{ij}(2)/\hat{h}_{ij}(3)]^{N_{ij}}$$

$$\chi^2 = \sum_{i,j} N_i \cdot [\hat{h}_{ij}(3) - \hat{h}_{ij}(2)]^2/\hat{h}_{ij}(2)$$

Hypotheses	Parameters	Dimension	DF
M2 vs M3	$\{m_{ij}(2)\}$	n-2	(n-1)(n-2)
	$\{m_{ij}(3)\}$	n(n-2)	

$$\lambda = \prod_{i,j} [\hat{m}_{ij}(2)/\hat{m}_{ij}(3)]^{N_{ij}}$$

$$\chi^2 = \sum_{i,j} N_i \cdot [\hat{m}_{ij}(3) - \hat{m}_{ij}(2)]^2/\hat{m}_{ij}(2)$$

Hypotheses	Parameters	Dimension	DF
H3 vs M3	$\{\varphi_i\}$	n	n(n-3)
	$\{m_{ij}(3)\}$	n(n-2)	

$$\lambda = \prod_{i,j} [\hat{h}_{ij}(3)/\hat{m}_{ij}(3)]^{N_{ij}}$$

$$\chi^2 = \sum_{i,j} N_i \cdot [\hat{m}_{ij}(3) - \hat{h}_{ij}(3)]^2/\hat{h}_{ij}(3)$$

APPENDIX F

TESTS OF THE MARKOV ORDER 1 AND 2 BURST MODELS

Hypothesis	Parameters	Dimension	DF	Notes
$M3_{S,1}$ vs $M3_{S,2}$	$\{p_{jk}\}$ $\{p_{ijk}\}$	$n(n-2)$ $n(n-1)(n-2)$	$n(n-2)^2$	$[N_{ijk}(g)]$: Markov 2 burst frequency matrix for group g . i,j,K = 1, ..., n member (numbered from highest participant to lowest) g = 1, ..., s group

$$\lambda = \prod_{i,j,k} [\hat{p}_{jk}/\hat{p}_{ijk}]^{N_{ijk}(\cdot)}$$

$$\chi^2 = \sum_{i,j,k} N_{ij\cdot}(\cdot)[\hat{p}_{ijk} - \hat{p}_{jk}]^2/\hat{p}_{jk}$$

$$\hat{p}_{jk} = N_{\cdot jk}(\cdot)/N_{\cdot j\cdot}(\cdot)$$

$$\hat{p}_{ijk} = N_{ijk}(\cdot)/N_{ij\cdot}(\cdot)$$

Hypothesis	Parameters	Dimension	DF
$M3_{S,1}$ vs $M3_{D,1}$	$\{p_{ij}\}$ $\{p_{ij}(g)\}$	$n(n-2)$ $n(n-2)s$	$n(n-2)(s-1)$

$$\lambda = \prod_{g,i,j} [\hat{p}_{ij}/\hat{p}_{ij}(g)]^{N_{ij}(g)} \qquad \hat{p}_{ij} = N_{ij}(\cdot)/N_{i\cdot}(\cdot)$$

$$\chi^2 = \sum_{g,i,j} N_{i\cdot}(g)[\hat{p}_{ij}(g) - \hat{p}_{ij}]^2/\hat{p}_{ij} \qquad \hat{p}_{ij}(g) = N_{ij}(g)/N_{i\cdot}(g)$$

Hypothesis	Parameters	Dimension	DF
$M3_{D,1}$ vs $M3_{D,2}$	$\{p_{jk}(g)\}$ $\{p_{ijk}(g)\}$	$n(n-2)s$ $n(n-1)(n-2)s$	$n(n-2)^2 s$

$$\lambda = \prod_{g,i,j,k} [\hat{p}_{jk}(g)/\hat{p}_{ijk}(g)]^{N_{ijk}(g)} \qquad \hat{p}_{jk}(g) = N_{\cdot jk}(g)/N_{\cdot j\cdot}(g)$$

$$\chi^2 = \sum_{g,i,j,k} N_{ij\cdot}(g)[\hat{p}_{ijk}(g) - \hat{p}_{jk}(g)]^2/\hat{p}_{jk}(g) \qquad \hat{p}_{ijk}(g) = N_{ijk}(g)/N_{\cdot j\cdot}(g)$$

Hypothesis	Parameters	Dimension	DF	Notes
$M3_{S,2}$ vs $M3_{,2}$	$\{p_{ijk}\}$	$n(n-1)(n-2)$	$n(n-1)(n-2)(s-1)$	
	$\{p_{ijk}(g)\}$	$n(n-1)(n-2)s$		

$$\lambda = \prod_{g,i,j,k} [\hat{p}_{ijk}/\hat{p}_{ijk}(g)]^{N_{ijk}(g)}$$

$$\chi^2 = \sum_{g,i,j,k} N_{ij\cdot}(g)[\hat{p}_{ijk}(g) - \hat{p}_{ijk}]^2/\hat{p}_{ijk}$$

APPENDIX G

FIXED VECTORS OF THE MARKOV ORDER 1 BURST MODELS

M1:

$$[x_1 x_2 x_3] \begin{bmatrix} 0 & p_2/(1-p_1) & p_3/(1-p_1) \\ p_1/(1-p_2) & 0 & p_3/(1-p_2) \\ p_1/(1-p_3) & p_2/(1-p_3) & 0 \end{bmatrix} = [x_1 x_2 x_3] \text{ ,}$$

$$p_1/(1-p_2)x_2 + p_1/(1-p_3)x_3,\ p_2/(1-p_1)x_1 + p_2/(1-p_3)x_3,\ p_3/(1-p_1)x_1$$

$$+ p_3/(1-p_2)x_2 = [x_1 x_2 x_3]$$

$$x_2/(1-p_2) + x_3/(1-p_3) = x_1/(1-p_1)$$

$$x_1/(1-p_1) + x_3/(1-p_3) = x_2/(1-p_2)$$

$$x_1/(1-p_1) + x_2/(1-p_2) = x_3/(1-p_3)$$

$$x_1(1/p_1 + 1/(1-p_1)) = x_2(1/p_2 + 1/(1-p_2)) = x_3(1/p_3$$

$$+ 1/(1-p_3))$$

$$x_1 + x_2 + x_3 = 1$$

$$\therefore\ x_k = p_k\overline{p_k}/c \text{ ;} \qquad \text{where} \quad c = (\Sigma p_k\overline{p_k})^{-1} \quad \text{and} \quad \overline{p_k} = 1 - p_k \text{ .}$$

H1: substitute for p_k .

M2:

$$[x_1 x_2 x_3] \begin{bmatrix} 0 & p_1 & p_2 \\ p_1 & 0 & p_2 \\ p_1 & p_2 & 0 \end{bmatrix} = [x_1 x_2 x_3]$$

$$[p_1x_2 + p_1x_3,\ p_1x_1 + p_2x_3,\ p_2x_1 + p_2x_2] = [x_1x_2x_3]$$

$$x_1 = (p_1p_2 + p_1)/(p_1^2 + p_2)\cdot x_2 = (p_1p_2 + p_1)/(p_1p_2 + p_2)\cdot x_3$$

$$x_1 + x_2 + x_3 = 1$$

$$[x_1x_2x_3] = [p_1(1 + p_2)/c,\ (p_1^2 + p_2)/c,\ (p_1 + 1)p_2/c]$$

where c is the appropriate normalizing constant.

H2: substitute for p_k .

M3:

$$[x_1x_2x_3]\begin{bmatrix} 0 & p_{11} & p_{12} \\ p_{21} & 0 & p_{22} \\ p_{31} & p_{32} & 0 \end{bmatrix} = [x_1x_2x_3]$$

$$[p_{21}x_2 + p_{31}x_3,\ p_{11}x_1 + p_{32}x_3,\ p_{12}x_1 + p_{22}x_2] = [x_1x_2x_3]$$

$$(p_{11}p_{31} + p_{32})x_1 = (p_{21}p_{32} + p_{31})x_2$$

$$(p_{12}p_{21} + p_{22})x_1 = (p_{21} + p_{22}p_{31})x_3$$

$$\frac{x_1}{(p_{21}p_{32} + p_{31})(p_{21} + p_{22}p_{31})} = \frac{x_2}{(p_{11}p_{31} + p_{32})(p_{21} + p_{22}p_{31})}$$

$$= \frac{x_3}{(p_{12}p_{21} + p_{22})(p_{31}p_{32} + p_{31})}$$

Simplify by using the fact that the row sums are all equal to one.

$$\frac{x_1}{(1 - p_{32}p_{22})} = \frac{x_2}{(1 - p_{31}p_{12})} = \frac{x_3}{(1 - p_{21}p_{11})}$$

$$[x_1 x_2 x_3] = [(1 - p_{22}p_{32})/c,\ (1 - p_{12}p_{31})/c,\ (1 - p_{11}p_{21})/c]$$

where c is the appropriate normalizing constant.

H3: substitute for p_{ij} .