
Robert Paul Sechrist
Louisiana State University and Agricultural & Mechanical College
INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*

4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

*For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.

UMI Dissertation Information Service

University Microfilms International
A Bell & Howell Information Company
300 N. Zeeb Road, Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Sechrist, Robert Paul

A COMPARATIVE ANALYSIS OF MULTI-DIMENSIONAL DIFFUSION MODELS: THE DIFFUSION OF THE PROHIBITION MOVEMENT IN THE UNITED STATES OF AMERICA, 1876-1919

The Louisiana State University and Agricultural and Mechanical Col. Ph.D. 1986

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages
2. Colored illustrations, paper or print
3. Photographs with dark background
4. Illustrations are poor copy
5. Pages with black marks, not original copy
6. Print shows through as there is text on both sides of page
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements
9. Tightly bound copy with print lost in spine
10. Computer printout pages with indistinct print
11. Page(s) lacking when material received, and not available from school or author.
12. Page(s) seem to be missing in numbering only as text follows.
13. Two pages numbered. Text follows.
14. Curling and wrinkled pages
15. Dissertation contains pages with print at a slant, filmed as received
16. Other

University Microfilms International
A COMPARATIVE ANALYSIS OF MULTI-DIMENSIONAL DIFFUSION MODELS: THE DIFFUSION OF THE PROHIBITION MOVEMENT IN THE UNITED STATES OF AMERICA, 1876-1919

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in The Department of Geography and Anthropology

by

Robert P. Sechrist
B.A., University of Pittsburgh, 1977
M.A., State University of New York at Binghamton, 1980
May 1986

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
ACKNOWLEDGEMENTS

I would like to thank several people who were of great assistance to me in the writing and analysis of this document. I owe a great deal to all the staff at Louisiana State University's Division of Research Services. They provided me with helpful assistance in terms of programming, methods, and content. The Division further supplied me with computer access when all other avenues of access were closed. I am indebted to Mr. Randy Hebert for teaching me the dBase language, and Dr. Marc Harris for his help in unraveling the mysteries of the political history of the late nineteenth century. The member of the Division's staff most deserving of my appreciation is, however, Alexandra Sparks. Never before, and probably never again, will I meet someone of her caliber and expertise.

I also thank the members of my committee for the efforts they put forth. Especially the considerable editorial work done by Dr. Milton Newton; work that converted this document from a meaningless jumble of quasi-English thoughts to what follows.

Also deserving my acknowledgement is the Inter-university Consortium for Political and Social Research. Accessing their census data bases kept me from almost a year of data entry.

Lastly, and most importantly, I thank my wife, Dr. Gail Sechrist. Gail has turned my life around and given me the impetus I needed to complete this degree. Without her, the completion of this task would be a hollow victory indeed.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENT PAGE**

**LIST OF TABLES**

**LIST OF FIGURES**

**ABSTRACT**

**INTRODUCTION**

**CHAPTER 1. HISTORY OF PROHIBITION IN THE UNITED STATES**

1.1 Before There Was Temperance
1.2 Temperance Organizations and Societies
1.3 Prohibition Prior to 1876
1.4 The Prohibition Movement from 1876 to 1919
1.5 Summary

**CHAPTER 2. THE DEVELOPMENT OF DIFFUSION RESEARCH IN GEOGRAPHY**

2.1 Origin of Diffusion Studies
2.2 Original Invention Versus Diffusion
2.3 Hagerstrand and Spatial Diffusion
2.4 The Network Diffusion Models of Anatol Rapoport
2.5 Recent Trends in Diffusion Studies
2.6 Summary

**CHAPTER 3. A DIFFUSION MODEL FOR THE PROHIBITION MOVEMENT**

3.1 The Adoption of Prohibition
3.2 Preparation of Input Data Dimensions
3.3 Nature and Organization of the Diffusion Model
3.4 Methods Employed in Analyzing and Interpreting the Results of the Diffusion Model
3.5 Summary

**CHAPTER 4. ANALYSIS OF SIMULATED DIFFUSION IN THE SPATIAL DIMENSIONS**

4.1 Analytical Prerequisites for Spatial Modeling
4.2 Associating Predictions and Reality
4.3 Analysis of the Distribution of Predictive Errors
4.4 Summary

**CHAPTER 5. ANALYSIS OF SIMULATED DIFFUSION IN THE URBAN-ETHNIC DIMENSIONS**

5.1 Building the Coordinate System
5.2 The Distribution of Counties in the Social Dimension
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Typical Record Generated by Program1 for Spatial Dimensions: The Relationship Between West Chester, Pennsylvania and the Twelve Nearest County Seats</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Accuracy of the Model in the Spatial Dimensions</td>
<td>83</td>
</tr>
<tr>
<td>4.2</td>
<td>Accuracy of the Model in the Spatial Dimensions in Selected Regions</td>
<td>84</td>
</tr>
<tr>
<td>5.1</td>
<td>Orthogonal Rotation of Factor Pattern and Standardized Scoring Coefficients</td>
<td>129</td>
</tr>
<tr>
<td>5.2</td>
<td>Selected Variables and Factor Scores for Selected Counties</td>
<td>133</td>
</tr>
<tr>
<td>5.3</td>
<td>Categorized County Factor Scores by Selected Regions</td>
<td>135</td>
</tr>
<tr>
<td>5.4</td>
<td>Accuracy of the Model in the Socio-Economic Dimensions</td>
<td>139</td>
</tr>
<tr>
<td>5.5</td>
<td>Accuracy of the Model in the Socio-Economic Dimensions, by Quadrant</td>
<td>144</td>
</tr>
<tr>
<td>5.6</td>
<td>Accuracy of the Socio-Economic Dimensions in 1884, by Quadrant</td>
<td>149</td>
</tr>
<tr>
<td>5.7</td>
<td>Accuracy of the Socio-Economic Dimensions in 1887, by Quadrant</td>
<td>152</td>
</tr>
<tr>
<td>5.8</td>
<td>Accuracy of the Socio-Economic Dimensions in 1892, by Quadrant</td>
<td>156</td>
</tr>
<tr>
<td>5.9</td>
<td>Accuracy of the Socio-Economic Dimensions in 1908, by Quadrant</td>
<td>159</td>
</tr>
<tr>
<td>5.10</td>
<td>Accuracy of the Socio-Economic Dimensions in 1909, by Quadrant</td>
<td>163</td>
</tr>
<tr>
<td>5.11</td>
<td>Accuracy of the Socio-Economic Dimensions in 1910, by Quadrant</td>
<td>164</td>
</tr>
<tr>
<td>5.12</td>
<td>Accuracy of the Socio-Economic Dimensions in 1916, by Quadrant</td>
<td>170</td>
</tr>
<tr>
<td>5.13</td>
<td>Accuracy of the Socio-Economic Dimensions in 1917, by Quadrant</td>
<td>171</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.14</td>
<td>Accuracy of the Socio-Economic Dimensions in 1918, by Quadrant</td>
<td>173</td>
</tr>
<tr>
<td>6.1</td>
<td>Accuracy of the 3-D Model Using Latitude, Longitude, and Urbanness by Regions and Urban Scores</td>
<td>185</td>
</tr>
<tr>
<td>6.2</td>
<td>Accuracy of the 3-D Model Using Latitude, Longitude, and Ethnic-Racial Composition by Regions and Ethnic-Racial Scores</td>
<td>192</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Ante-Bellum Local Option Legislation by State</td>
<td>14</td>
</tr>
<tr>
<td>1.2</td>
<td>First Year of State-wide Ante-Bellum Prohibition by State</td>
<td>16</td>
</tr>
<tr>
<td>1.3</td>
<td>Number of Dry Counties by Year in the United States</td>
<td>18</td>
</tr>
<tr>
<td>1.4</td>
<td>Percent of Counties Dry by Year in the United States</td>
<td>18</td>
</tr>
<tr>
<td>1.5</td>
<td>Year of Post Civil War Passage of General Local Option Legislation</td>
<td>21</td>
</tr>
<tr>
<td>1.6</td>
<td>First Continuous Year of Prohibition Prior to 1920 by State</td>
<td>27</td>
</tr>
<tr>
<td>2.1</td>
<td>Accumulated Intervals and Probability for Mean Information Field</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>Flowchart of Program1, Data Structure Creation</td>
<td>57</td>
</tr>
<tr>
<td>3.2</td>
<td>Spatial Distribution of County Seats About Chester County, Pennsylvania</td>
<td>61</td>
</tr>
<tr>
<td>3.3</td>
<td>Flowchart of Program2, The Simulation Model</td>
<td>65</td>
</tr>
<tr>
<td>3.4</td>
<td>Flowchart of Program3, Merging-Omega File Creation</td>
<td>72</td>
</tr>
<tr>
<td>4.1</td>
<td>Results of Diffusion in the Spatial Dimensions, 1878</td>
<td>87</td>
</tr>
<tr>
<td>4.2</td>
<td>Results of Diffusion in the Spatial Dimensions, 1881</td>
<td>91</td>
</tr>
<tr>
<td>4.3</td>
<td>Results of Diffusion in the Spatial Dimensions, 1885</td>
<td>94</td>
</tr>
<tr>
<td>4.4</td>
<td>Results of Diffusion in the Spatial Dimensions, 1887</td>
<td>95</td>
</tr>
<tr>
<td>4.5</td>
<td>Results of Diffusion in the Spatial Dimensions, 1891</td>
<td>96</td>
</tr>
<tr>
<td>4.6</td>
<td>Results of Diffusion in the Spatial Dimensions, 1904</td>
<td>102</td>
</tr>
<tr>
<td>4.7</td>
<td>Results of Diffusion in the Spatial Dimensions, 1907</td>
<td>107</td>
</tr>
<tr>
<td>4.8</td>
<td>Results of Diffusion in the Spatial Dimensions, 1908</td>
<td>108</td>
</tr>
<tr>
<td>4.9</td>
<td>Results of Diffusion in the Spatial Dimensions, 1909</td>
<td>110</td>
</tr>
<tr>
<td>4.10</td>
<td>Results of Diffusion in the Spatial Dimensions, 1910</td>
<td>111</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.11</td>
<td>Results of Diffusion in the Spatial Dimensions, 1916</td>
<td>115</td>
</tr>
<tr>
<td>4.12</td>
<td>Results of Diffusion in the Spatial Dimensions, 1917</td>
<td>116</td>
</tr>
<tr>
<td>4.13</td>
<td>Results of Diffusion in the Spatial Dimensions, 1918</td>
<td>118</td>
</tr>
<tr>
<td>5.1</td>
<td>Distribution of Wet and Dry Counties in the Socio-Economic Dimensions, 1876</td>
<td>141</td>
</tr>
</tbody>
</table>
ABSTRACT

Geographical study of the spatial spread of innovations has revealed much about the spatial components of innovation diffusion, but innovations diffuse in other dimensions as well. This study compares the ability of the classic spatial diffusion model with diffusion modeled in constructed socio-economic dimensions. Each model is compared with its counterpart in its ability to simulate and explain the growth of the prohibition movement during the period 1876 to 1920.

The various models indicate that prohibition spread in several dimensions simultaneously. The spatial diffusion model demonstrates that the innovation spread first to the West and then back North from a Southern core. The socio-economic model indicates that prohibition spread from counties with high black and Protestant concentrations to counties whose residents were proportionally more white and Catholic. Evaluation of the urban hierarchical model revealed that prohibition spread from more rural to urban counties. A spatial diffusion model using county ethnic-racial composition reinterpreted and clarified previous findings.

The comparison of the models led to the realization that spatial diffusion models, while explaining a large proportion of the variance observed, are inadequate. A model that encompasses the special cases and ad hoc fixes necessary in spatial diffusion modeling was required. The concept of general n-dimensional models is offered as a solution to these problems with spatial diffusion models. Spatial diffusion models are reclassified as a specific type of a general two-dimen-
sional model. The special case of the spatial model, urban hier-
archical diffusion, is similarly reclassed as but one type of a gen-
eral three-dimensional model. Two and three-dimensional models are
also only special cases of n-dimensional diffusion models. Methods
and guidelines for creating and implementing n-dimensional models are
presented in the final chapter.
INTRODUCTION

In this study, a spatial diffusion model is compared with a social diffusion model. This comparison is performed to determine if a diffusion model based on propinquity provides a better explanatory framework than a model based on similarity. Out of this comparison comes the basis for the construction of diffusion models based in more than two dimensions. The diffusion of the prohibition movement in the United States is used as the vehicle for the discussion and models. Diffusion studied in the spatial dimensions has long been a major focus of geographical inquiry, but diffusion can be studied just as profitably in the social dimensions. Hagerstrand identified the potential for studying diffusion by using either approach when he said, "The diffusion of an innovation propagates in two dimensions, the spatial and the social" (Hagerstrand 1952, 356). At that time, Hagerstrand chose to examine the spatial dimensions of innovation diffusion. The time has come to examine diffusion in the social dimensions.

Rather than create only a diffusion model based on social characteristics of adopters, the research reported here generated both spatial and social versions of a diffusion model modified from Hagerstrand. This strategy allows the use of one of the most powerful tools in scientific research, the comparative method. Each version's ability to predict prohibition status is evaluated, measured, and compared to the other, enabling the researcher to discern what each of the two versions of the same model do and do not tell us about adopt-
tion behavior, the prohibition movement, and patterns of information flow.

The model used to simulate both spatial and social diffusion is, in principle, the same model that has been used by students of spatial diffusion for the past thirty years. In each version of the model, adopters disperse information regarding an innovation to those who have not yet adopted (potential adopters). In the spatial model, information flows between neighbors, while in the social model information flows between individuals who have the greatest socio-economic similarity. The two versions of the model produce a surprisingly similar degree of accuracy, but the distributions of predictive errors in each version differ greatly. Each version of the model provides a differently skewed explanation of the events.

Prohibition, the legal banning of manufacture, sale, possession, and consumption of alcoholic beverages, was selected as the innovation of study for several reasons. In the majority of states, each county independently decided its own prohibition status. For the most part, this decision was accomplished by local option elections in which the electorate of a county voted either for or against prohibition. The counties of the forty-eight coterminous states provide a large finite universe of observational units. At any given time during the study period, 1876-1919, each observational unit was either wet (not a prohibition county) or dry (a county where prohibition was the law).

Information necessary for the implementation of the spatial model is readily available for the 3030 counties that existed in 1920. The spatial model requires a set of fixed coordinates for each county, and the latitude and longitude of the county seat filled this role. The
social model requires appropriate social and economic data for each county. From the 1920, 1910, and 1900 Censuses of Population, the following data were selected: total population, percentage foreign born, percentage Negro, percentage who worked in industry, and percentage Roman Catholic. From the census of 1910 three additional variables are included: population density in 1910, percentage urban in 1910, and percentage in urban 1900. The percent Catholic is interpolated from the Censuses of Religious Bodies of 1896, 1906, 1916, and 1926. Because census data were used from as far back as 1896, only counties that existed in 1896 were included in the social model runs (n=2499). The prohibition status of each county during the period 1876-1919 was gathered from a wide variety of archival sources. Histories of the prohibition movement, although available for many states, often failed to specify the prohibition status of individual counties. Overall, however, they tend to be the best sources of information. After two years of data collection, it was possible to compile a data base consisting of the prohibition status of all counties in the continental United States.

Other reasons for selecting prohibition as the innovation for study include: 1) a deep and abiding personal interest in the growth of mass movements in general and in prohibition in particular; 2) a desire to see in cultural geography an emergence of diffusion research on ideological innovations; 3) the need to study a well documented social movement that has run its course; and 4) the hope of being able to explain why adoption occurs in terms other than those offered by economic geographers who study the diffusion of technological innovations and whose explanation for adoption ultimately seems to rest on
the notion that new is desirable.

The study period, 1876-1919, best coincides with the growth of the movement. There were, of course, dry counties prior to 1876, but the mid 1870s saw the rebirth of prohibition sentiment. Three separate political events heralded this rebirth: the formation in 1872 of the Prohibition Party, the founding of the Women's Christian Temperance Union (WCTU) in 1874, and the enactment of local option laws in several Southern states. Southern religious groups also first spoke out against the evils of drink during this decade. Further, the reliability of accounts of prohibition status of counties increased dramatically in that decade. Finally, the passage of a decade following the Civil War allowed the country to return to normal operation of civil law. The year 1919 terminates the study because national prohibition was enacted that year.

Chapter one discusses the historic background and events of the prohibition movement and further clarifies the choice of the study period. Chapter two deals with the historical development of spatial diffusion models as they are reflected in current research efforts. Chapter two also provides a discussion of the elements of spatial diffusion and their interpretation.

Chapters three, four, and five form the body of the study. Chapter three describes the diffusion model and the analytical techniques used in this study. Chapter four describes the analysis of diffusion modeled in the spatial dimensions. Chapter five, after providing justification for the implementation of the model on the social dimensions, describes and analyzes its implementation and results in the social dimensions in the same terms as with the spatial
model in chapter three.

In chapter six, a model is created that attempts to merge the best features of the model variants discussed in chapters four and five. This model operates with three variables and is therefore termed three-dimensional. This model is implemented on two sets of three dimensions, both of which include the spatial dimensions and one of the two social-economic dimensions as the third variable. The inclusion of a third dimension in the creation of the data structure redefines the proximity of counties to each other. The three-dimensional model is found to provide a more detailed explanation of why potential adopters adopt, while maintaining the level of model accuracy. The inclusion of a third dimension does not necessarily improve accuracy and may reduce it compared to a two-dimensional model using two of the same variables. Three-dimensional modeling also defines a method for developing and defining an n-dimensional model that can account for all variance.

Finally, new insights derived from the diffusion analysis regarding the growth of prohibition are presented as well as a summary of the advancements made in simulation diffusion models themselves.
CHAPTER ONE

HISTORY OF PROHIBITION IN THE UNITED STATES

1.1 Before There Was Temperance

The colonial period was an era of marked intemperance. People believed strong drink to be the greatest gift God had given to man. Liquor eased the colonists through the pain of life and gave them the will to continue in a hard world. It was therefore a very common practice to start the day with a drink and to drink whenever the occasion presented itself. During early colonial times, perhaps the most intemperate group of people were the clergy (Asbury 1950, 13). When he dropped by to visit a member of his congregation, a minister was met with a drink, and only the rudest of congregations would reject the offer, regardless of his condition (Fehlandt 1904).

Hard liquor had the endorsement of doctors, who prescribed it for practically every affliction from painful teething in infancy to the aches of old age. According to the medical mythology of the era, rum-and-milk was a boon to pregnant women, as well as nursing mothers. Rum-soaked cherries supposedly prevented colds. The plethora of cure-all tinctures, tonics and elixirs contained mainly alcohol and colored water. A life insurance company increased its premiums by 10 percent for the abstainer, whom it considered 'thin and watery, and as mentally cranked, in that he repudiated the good creatures of God as found alcoholic drinks' (Kobler 1973, 26).

Alcohol served extensively as a medicinal agent, but heavy drinking was believed to be its own reward in many other ways. People expressed their joys and sorrows through their drinking. Great quantities were consumed at funerals, weddings, and ordinations alike. Inhiting heavily in liquors seems to have been the preferred way to survive the boredom of winter and the rigors of summer.
The laws controlling liquor at that time reflect the American colonist's feelings about liquor. The proprietor of an inn was commonly required to keep spirits on hand for weary travelers and local customers. The liquor dealer was one of the major sources of public revenue, as most colonies required the innkeeper to purchase a liquor license, the proceeds from which went to the public coffers. Further, in many a village, the tavern was the only public structure and was as a result the meeting hall, courtroom, and in some instances the church (Asbury 1950, 24). Otherwise, sellers felt little legal control. The drinker had the responsibility to maintain good behavior during and after imbibing.

On an individual level, drunkeness to the point of stupor was deplored, while drinking consistently and steadily was encouraged. It was a common practice to single out those who were known to imbibe too heartily and to deny them drink in the future, fine them, whip them, or publically humiliate them (Cherrington 1920a; Chidsey 1969).

There were, of course, during colonial times a few individuals who spoke out against strong drink, but none spoke out against all drink. The first well documented instance of an individual's speaking against excessive drinking in the colonies occurred in 1726 when "Reverend Cotton Mather, D.D., of Massachusetts colony together with 22 other ministers published 'A Serious Address to Those Who Unnecessarily Frequent the Tavern'" (Cherrington 1920a, 33). The good ministers did not, however, chastise those who drank at home, or at parties, but only those who "Unnecessarily Frequent the Tavern."

The first prohibitory action taken in North America came in 1737
when Governor William Oglethorpe of Georgia banned the importation of spirits into his colony. He felt that his action was necessary to obtain a proper level of work from his conscripted colonists. The ban did not last long; in only five years, it was lifted when the crisis of founding the new colony was over (Asbury 1950, 21). In any case, it failed dismally, just as national prohibition would 200 years later.

The origins of a temperance movement lie, however, after the close of the Revolution when in 1785 Dr. Benjamin Rush published a series of articles on "The Effects of Ardent Spirits Upon the Human Body and Mind." These works, from the pen of the eminent Pennsylvania physician and member of the Continental Congress, stirred public anxiety regarding the dangers of excessive drinking. In particular, Dr. Rush's "An Essay on the Effect of Ardent Spirits upon the Human Constitution" inspired the first temperance organization (Dorchester 1888, 171).

1.2 Temperance Organizations and Societies

Before Dr. Rush wrote his articles, and for some twenty years after, there was no temperance movement. There was no movement because there was no organization of temperance advocates, merely scattered voices in a sea of inebriety. The first temperance organization was founded in 1813 by Dr. Billy J. Clark of Saratoga, New York. Dr. Clark, having read Dr. Rush's writings, founded a society which had as goals for its members abstinence from strong liquor. A key feature of the Union Temperance Society of Moreau and Northumberland, as Dr. Clark named his organization, was a signed pledge in which members promised not to drink liquor, except under orders of a physician, under a penalty
of a twenty-five cent fine (Dorchester 1884; Cherrington 1920a; Kobler 1973).

During the next twenty years, myriad local temperance societies were established. By 1820, it seemed that, each community had its own temperance society, composed of sober citizens totally unaffiliated with temperance societies in neighboring communities. The number of these societies is very difficult to estimate as is the number of members in each; they were often short-lived and died ignominiously. Societies were established with great fanfare, and every time a new society was founded, members abandoned the former society and joined with the new. That pattern plagued the movement throughout the 1800s.

In the late 1820s, the community-based societies gave way to state-wide organizations. The Massachusetts Society for the Suppression of Intemperance, organized on February 12, 1813, in Boston, was the first state-wide temperance organization (Cherrington 1920a, 91). The rapid adoption of state organizations in the New England and the Middle Atlantic regions indicates how ready these states were for the hierarchical leap. In 1829, there were already eleven state temperance societies embracing more than a thousand local chapters (Gordon 1932, 7). These societies were most densely concentrated in the Northeast, primarily in New England and New York, but there were outposts of temperance activities in Mississippi and Georgia both of whose state temperance societies were established before 1833. The coordinating role of state-level temperance societies was brought to an end by nation-wide temperance organizations that, while first appearing in the late 1820s, did not take complete control of state and local societies until the 1840s.
The first of these national organizations, The American Society for the Promotion of Temperance, was founded on February 13, 1826, in Boston (Dorchester 1884; Cherrington 1920a). This organization grew rapidly at the expense of local and state organizations. Its name was changed in 1832 to the American Temperance Society. It remained a major influence throughout the nineteenth century. Starting in 1842, however, a series of new organizations challenged The American Temperance Society as coordinator and controller of temperance sentiment.

The Order of the Sons of Temperance, organized in New York, and The Independent Order of Rechabites, introduced from England, were both organized in 1841. In 1845, The Order of the Templars of Honor and Temperance was founded. The year 1846 saw the organization of the Cadets of Temperance in Germantown, Pennsylvania, and the birth of the Order of Good Samaritans in New York (Cherrington 1920a, 94).

These new organizations shared several basic similarities among their goals and organizational principles. First, starting in 1836 with the American Temperance Society, they all insisted that members abstain from liquors (in later years these societies tended to insist on abstainence from all alcoholic beverages) (Asbury 1950, 36). Second, the intent of the members to abstain was sanctified by the signing of a pledge. Third, the membership rosters of individual societies were very unstable. An individual might have joined several of these organizations during the decade, abandoning his active membership in the older, no longer popular organization for membership in a new, growing society. When a new society was organized it would sweep the country (the densest concentrations of chapters of any society were to be found in the Northeast, but chapters could be found throughout the
nation) forming local chapters in many cities and communities. An older, less vibrant society could easily die out. Fourth, the most successful and long lasting of these organizations were secret societies, complete with handshakes, passwords, and dues (Kobler 1973, 71-75). Fifth, the societies were for temperate men; society members generally maintained that the inebriate was beyond help, and the societies therefore sought only members who did not drink excessively. Sixth, these organizations were for men only (Daniels 1878, 195-211; Cherrington 1920a, 95-134).

There were, of course, exceptions to the general patterns. Two societies of exceptional nature deserve mention. The first, The Cadets of Temperance, founded in 1846 and directly linked to the Sons of Temperance, was exceptional because it organized boys and young men to keep them on the right path. The second exceptional society was the Washingtonians, a society for reformed drunkards, that was founded on April 2, 1840. The Washingtonian Society was formed by six drinking buddies who reformed simultaneously after hearing a temperance lecture and who spread their reformation to many others across the nation.

Who were the people who joined these societies? What were their individual goals? The movement to establish societies was assimilative reform initiated by the middle class (Gusfield 1963, 29). The core of any local chapter belonged to the community's middle class. "The middle class advocated abstinence as a distinct symbol of its way of life and offered temperance to the lower class as a means by which its members could raise themselves to a higher level of economic well being" (Nelson 1968, 38). Membership in temperance societies indicated status in the community. Members of the lower class who aspired to better things
joined the movement to associate themselves with the middle class and to place themselves where economic opportunity existed. Men who participated in temperance activities were those who had been assimilated into the mainstream of American life (Gusfield 1963, 71).

All the organizations established between 1813 and 1850 shared the overriding similarity that they were all based on the premise that Demon Rum could be eradicated if only men would put their minds to this great task and control their base desires. Temperance men believed that, with group support and logical arguments on the dangers and effects of liquor, they could alter the behavior of drunkards and social drinkers. As a group, temperance leaders failed to foresee the individual who would neither succumb to the arguments of moral suasion nor identify with success as conceived by the middle class. As temperance workers became frustrated in the failure of their method of moral suasion, they took a stronger stand and became prohibitionists.

1.3 Prohibition Prior to 1876

Easton, Massachusetts, in 1830 was, the first community to ban the sale of liquor (Clark 1888). It was not the last. In 1831, Plymouth County, Massachusetts, became the first county in the nation to go dry. By 1840, nine of the fourteen Massachusetts counties were dry (Tyrrell 1979; Hampel 1982, 171). A weak inverse correlation was discovered between county population density and the first year the county was dry (r=-.4233) among Massachusetts counties 1831 to 1852. This correlation persistently recurs during the diffusion of prohibition in the United States, demonstrating that the movement's primary appeal lay in rural areas.
Massachusetts was not the only state to have dry counties. The state of Georgia, in 1833, enacted a law that allowed the rural, coastal counties of Liberty and Camden to hold referenda, henceforth referred to as local option elections, on whether it would be legal for liquor to be sold in those counties. Both counties elected to become dry territories, and for several years after the elections, no one could legally obtain a drink in those counties (Dunn 1877; Scomp 1888). New York enacted its first local option law in 1845, and at the local option elections of 1846, 728 of 856 towns and cities voted out the saloon (Asbury 1950). Legislation that enabled local option elections passed in twenty states prior to the Civil War. Two general paths guided the spread of local option legislation in ante-bellum America (Fig. 1.1). The first carried it south and west from the New England hearth, while the second led from its Georgia-Alabama core slowly westward across the Deep South.

The local option movement, like the moral suasion movement before it, was soon discovered to be inadequate. Persons desiring a drink could simply travel to a wet town, perhaps less than a day's ride away and stock up, or he could let friends and business associates make the trip for him. The first temperance leader to address this problem was Neal Dow of Maine. Dow became motivated in temperance activities when a bartender would not stop serving a customer who, in Dow's opinion, had had too much (Dow 1898). Under Dow's direction, the city of Portland, Maine, went dry in 1843. With this success under his belt, Dow started a campaign for state-wide prohibition which succeeded in 1851. "There was no nonsense about the Maine Law, and it contained no loopholes through which violators might wriggle. It prohibited the
Figure 1.1

Ante-Bellum Local Option Legislation by State

Legend
- 1833-1837*
- 1838-1839
- 1840-1841
- 1845-1847
- 1848-1854
- No Passage

*Includes Massachusetts which had dry country by 1831, but did not pass local option until 1845.
sale, the keeping for sale, and the manufacture of all intoxicating liquors" (Asbury 1950, 59). This prohibitory law remained in effect until 1933, except for a brief repeal in 1857.

Most of the other Northern states soon followed the lead set by Maine (Fig. 1.2). By 1855, over 600 counties in the United States had gone dry (Fig. 1.3). In 1855, practically the entire northern half of the country was dry. Pennsylvania, Wisconsin, New Jersey, and North Carolina had all held referenda that were narrowly won by the wets. Most of the elections won by the drys were quickly reversed; New York state was dry for less than one year, and by 1860, only Maine, Vermont, New Hampshire, Rhode Island, Michigan, Iowa, Ohio, and Massachusetts were still dry.

An explanation of why states went dry in the 1850s goes deeper than simply the failings of moral suasion in its attempt to defeat Demon Rum. The ante-bellum prohibition wave was also closely tied to the Know-Nothing political party and nativistic sentiment in general. The Know-Nothing Party urged stopping future immigration from Europe especially from countries other than England, and nativists held that recent immigrants should be treated as second class citizens or, better yet, as non-citizens. Both groups believed that America was only for people born in America.

With the rapid influx of heavy drinking immigrants from Ireland and Germany in the late 1840s and 1850s, the gains made through temperance legislation and assimilative actions of the middle class were threatened; these new Americans were prone neither to abstaination nor to voting for prohibitory laws. This fact scared prohibition leaders into bed with nativistic leaders; as similar threats would scare
future prohibition leaders into coalitions with the leaders of other hysteria-driven movements. The coalition had influence sufficient to obtain the passage of the state prohibition laws. As the new immigrants obtained citizenship and became politically active, however, many of the laws were repealed, most within a year or two of passage.

The immigrant groups acting alone did not bring about the repeal of the laws. The people who made and sold liquor provided strong support for repeal activities and legislation. The coalition of prohibition and nativism broke up after the passage of the laws, and the prohibitionist either went home satisfied that he had saved his fellow man, or he immersed himself in fighting the other great putative evil of the time, slavery. As a result, few actively opposed the repeal of the recently passed prohibitory laws.

In the South, with few immigrants and no nativistic movement, a Southern prohibition-nativistic coalition never formed, killing the chance for politically motivated ante-bellum prohibition laws in the South. More inhibiting to the development of a strong Southern prohibition movement was the fact that Southern churches refused to become involved in civil matters. Southern churches originally took this stance because of the slavery issue, "considerations of consistency would [sic] shut the Church [Methodist Episcopal, South] out from exercising any influence on 'the liquor question ...'" (Farish 1938, 98).

Interest in prohibition declined rapidly after 1856. It was the last year when any state went dry, and the beginning of the repeal activities of the liquor interests and immigrants. The prohibitionists had gone home, triumphant, after the passage of the prohi-
Figure 1.3

Figure 1.4
bitory laws. Had they kept active, their organizations they might have succeeded against the wave of repeals, but by that time national attention had focused on a new cause, the impending Civil War (Asbury 1950, 61; Gallaher 1933, 220).

1.4 The Prohibition Movement from 1876 to 1919

While the nation remained immersed in reconstruction, the prohibition movement, in 1872, again bid for popular favor. Renewed interest became tangible in the creation of the Prohibition party. The first Prohibition presidential candidate, James Black of Pennsylvania, received, in the election of 1872, a mere 5602 votes (Cherrington 1920a, 188). Although the Prohibition party grew in power during the late nineteenth century, it was unable to elect candidates to major offices. The death of the Prohibition party was the direct result of the major parties' coming to favor prohibition in the early 1900s (Colvin 1926).

The second event indicating a resurgence of prohibition sentiment occurred with the founding of the Women's Christian Temperance Union (WCTU) on November 17, 1874. The WCTU was unique because it was the first organization of women, for women, and by women. The members of the WCTU quite flamboyantly stated and acted to achieve their goals. The WCTU organized women to march at the polls where local option elections were held, and they created and provided alternate, wholesome forms of entertainment for young people. The WCTU grew and flourished during this period. As the only organization that American women had, it served as the focal point for not only prohibition activity, but later for suffrage activities (Beard 1962; Duniway
Until about 1895, these two organizations served as the principal vehicles for anti-liquor sentiment. The older societies persisted, but merely as social organizations; they no longer actively sought to obtain passage of prohibition laws. The Prohibition party put up candidates for practically every office, but they won very few posts. The WCTU marched, sang, and pleaded with drunkards and elected some officials (indirectly), but made little headway in reaching their goals of prohibition for their communities, counties, states, and nation. Neither group had systematic success (Fig. 1.3 and 1.4). The period from 1876 to 1895 was a one of very slow growth of prohibition territory.

These groups did, however, have some successes. Many state legislatures, influenced by their activities, reenacted general local option legislation (Fig. 1.5). Southern states tended to take the lead in enacting this type of legislation late in the nineteenth century. More perhaps as a result of actions by the clergy than any other group or lobby efforts. The Northeast and Midwest states did not reinstate local option legislation until after 1900. Maine and Kansas, which went dry in 1881, however, never had local option because they were dry territory. Vermont and New Hampshire were not local option states before 1904 because they were, until that time, under state-wide prohibition. The South, during the ante-bellum period, did not adopt complete state-wide prohibition (Fig. 1.2), and therefore was perhaps not as leary of passage as the North and West, which were much slower to readopt a strategy which had previously failed.

Four factors correlate with the adoption of local option legisla-
Figure 1.5

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
tion in the South prior to other regions. First, the South was primarily a rural region, while the North was becoming urban, and the West was untamed. Second, the South had few immigrants, while the North and West were being inundated with hard drinking Germans and such predominantly Catholic groups as the Irish and later the Italians (Campbell 1977; Ward 1971, 71-81). Third, the actions taken by the Southern states seem to have been linked to the wave of religious fundamentalist fervor that passed through the South in the 1870s and 1880s (Whitner 1945, 19; Pearson and Hendricks 1967, 60). Southern church leaders, who could not speak on civil matters before the war because of slavery, were permitted to speak out against liquor and its abuses after the war (Farish 1938 308-14). Fourth, the rise of the Populist party, which was present only in the South, brought about the spread of prohibition by two means.

First, the Populists, who were both white and black and stood for racial harmony, were supporters of prohibition. Second, Southern conservatives, argued that blacks, being incapable of holding their liquor, would sell their votes for drinks and therefore should not be permitted to vote (on liquor issues at the least). This argument was primarily made in an attempt to quash the Populist movement by disenfranchising Negroes. The Southern conservatives were not necessarily against prohibition and may have used prohibition at the county level to control blacks and poor whites within their counties (Richland Beacon 1885; Walton and Taylor 1969; Walton 1972; Woodward 1974, 60-64). These factors combined to make the South after the Civil War the seed region for prohibition because, as the analysis of diffusion in the socio-economic dimension reveals, the places most likely to adopt prohibition had rural characteristics, a small number of foreign born residents, a small
percentage of Catholics, a high percentage of Negro residents, and local Protestant churches that gave strong support to the movement.

The states on the Great Plains approached the liquor problem from another direction. Entire states adopted simultaneously. Iowa and the Indian Territory were already dry in 1876. Kansas adopted in 1881. North and South Dakota adopted in 1889, but South Dakota repealed in 1897. The Indian Territory as Oklahoma adopted prohibition as a part of its constitution in 1907.

With the close of the 1880s, the prohibition movement entered into what appeared to be a period of decline and stagnation as evidenced by the reduction of the number of dry counties in the United States (Fig. 1.3). During the 1890s, few new prohibition laws were passed, and the remaining prohibition territory in the North was drastically reduced as both Ohio and Iowa returned into the wet column. Many counties in states with local option alternated between wet and dry status, some on an annual basis. This decline in effective prohibition activity is associated with several events: the capture of public and media attention by the Spanish American War, the depression of 1896 (Clark 1933), and the Supreme Court's devastating decision to uphold the original package law in 1893.

Economic depressions during the nineteenth century had a profound effect on the prohibition movement. Generally, public sentiment on moral issues reversed itself during a depression. Before the depression of 1896, for example, the movement was gathering strength, but this trend reversed after the depression of 1896. An individual's stance on prohibition changed with economic dishevels because he was seeking solace for his plight. At the group level, this abandonment, or aquisi-
tion, of morals is observed as general shifts in public opinion on moral and religious issues (Clark 1933).

Coming on the tail of the depression of 1896, the Spanish-American War kept public attention on matters other than prohibition. On the home front, newspapers had no room for prohibition stories, nor did the people have time to talk of it when they could discuss battles. Many wholesome young men went off to war, and in the company of other young men, took up habits that they might not have acquired at home. When these men dispersed after the war, their war experiences and new found vices brought them out against prohibition.

The Original Package Law permitted liquor to be sold in prohibition territory as long as it was shipped and sold in its original package (the package in which it had been shipped across the state line). The Court held that local prohibition of liquors shipped in the original package was an attempt by local government to regulate interstate commerce, a power reserved to Congress. As a result, original package shops sprang up throughout dry territory and could not be closed by local laws or wishes. Even though liquor dispensed at original package shops could be sold for take-out purposes only, the damage was done. Demon run slipped back into dry communities throughout the nation. The wind had been taken out of the prohibitionist sails (Columbia Law Review 1919, 114).

The organization that put the wind back in the sails of the prohibition movement was the Anti-Saloon League (ASL), which was founded in 1893 at Oberlin, Ohio. The ASL received wide support from the older partisan organizations, and it grew rapidly (Blocker 1976, 141). Unlike all previous organizations, the ASL did not represent any specific group
of Americans, and the organization became strong enough not to have to form coalitions with other political movements. The ASL had much greater, and more denominationally diverse, clerical support than any previous organization. The ASL backed any candidate, regardless of party, who favored prohibition and had a proven record. This tactic of single-issue endorsement eventually filled the state legislatures and Congress with sympathetic men and brought about substantial territorial gains.

Democratic and Republican candidates were content to sit on the fence on this issue; when prohibition did emerge as a campaign issue both candidates would oppose licensing the sale of liquor. People did not join the ASL as they had joined the secret or social societies; they joined the ASL at the ballot box when they voted against liquor. Further, the ASL had both male and female officers. A large percentage of their officers were clergymen representing many denominations. With this broad base of support, they were able, at first, to organize local residents to vote saloons out of their communities (Odegard 1928).

The Anti-Saloon League followed a strategy of consolidating their gains by moving up the hierarchy of political units. So, after a majority of a county's communities were dry, they next tried to dry up the county (in states where the county was the lowest level of local option, they of course started with the county). Then, after most of a state's counties were dry, they began to work for state-wide prohibition. In fact, the ASL regarded local option as only a "temporary expedient in those states where statewide prohibition was hopeless for a long time to come" (Colvin 1926, 359). They must certainly have viewed
state-level prohibition in the same light when they thought of national prohibition. There even exist records of the ASL's being against certain state-wide prohibition referenda and activities. "In Missouri, for example, league leaders claimed to have advised against the initiative petition for a vote on statewide prohibition, although they joined the campaign once the necessary signatures had been secured" (Blocker 1976, 216). The ASL opposed the move because league leaders knew it would fail, and presumably they felt that a single loss was more detrimental to their cause than could be compensated for by several victories. The effectiveness of the ASL did not really begin to be felt in most states before 1905.

The first states to become dry territory as a result of the action of the Anti-Saloon League were in the South. In these states, the ASL teamed with fundamentalist religion to produce the nation's first dry region since the North had dried up during the mid 1850s. The first Southern state to go dry was Georgia, in 1908, which was followed in 1909 by Tennessee, North Carolina, Mississippi, and Alabama (Odegard 1928). Alabama repealed state-wide prohibition in 1912, reverting to county local option, but reenacted it in 1915.

From 1909 on, the ASL was able to bring the full force of its national organization to bear upon each state's decision makers, whether they were voters in referenda or the state legislators. By 1910, there were eight dry states, five of which could be credited to ASL action. From 1910 on, the number of dry states increased steadily (Fig. 1.6). As in local option legislation, the South took the lead, most states going dry before 1910. The Western states followed, about half of which were dry before 1918. The Northeastern and
Figure 1.6

First Continuous Year of Prohibition Prior to 1920 by State
Midwestern states, except Maine and Michigan, failed to enact state-wide prohibition prior to National Constitutional Prohibition, which became effective on January 1, 1920 (Odegard 1928).

1.5 Summary

The diffusion of prohibition was a long, slow, involved process. Distinct spatial and social trends appear throughout the history of the movement. In the early years, the movement, which had originated in New England, diffused into the rest of the Northeast and from there into the Midwest. After adoption and rejection in these regions, the movement was relocated and revitalized in the South and Plains, from whence it spread into the West and then to the remainder of the nation.

The movement started among the clergy, physicians, and upright citizens. In the early years the movement provided a mechanism for social interaction and assimilation of the lower classes and immigrants into the middle class. The goals of the movement later became a weapon that was used against the immigrants by the Know-Nothing party and other nativists, against Catholics by fundamentalist Protestants, against Negroes by whites, and against the growing urban population by rural inhabitants. Each of these groups at one time or another pointed at their opposite and declared them both drunkards and the cause of ruination in America.

The prohibition movement became a growing concern whose diffusion can be readily traced after the three groups: the Prohibition party, the Women's Christian Temperance Union, and the Anti-Saloon League took the helm and coordinated the movement's dispersal. After
1876, the South was the growth pole for the movement at the same time it was waning in the North. Southern states were the first to institute local option legislation after the Civil War. Plains states were the first to initiate state-wide prohibition during the study period. From the Southern and Plains core prohibition territory spread to the Western and Northern states, starting about 1905. Prior to the enactment of National Constitutional Prohibition in 1920 there were 24 dry states, most of which were located in the South and West.
CHAPTER TWO

THE DEVELOPMENT OF DIFFUSION RESEARCH IN GEOGRAPHY

Like the preceding chapter, this one recapitulates historic events. Chapter one told of the events which led to national prohibition. Chapter two, however, discusses the history of models and theories used to explain and follow the growth patterns of innovations. The following discussion is a highly selective history that is intended to lead the reader from point to point in an attempt to familiarize him with the status of diffusion research as the author sees it. Both the preceding chapter, and this one, set the stage for the remainder of the dissertation where the diffusion of prohibition is discussed in terms of propinquity and similarity models.

2.1 Origin of Diffusion Studies

In the early years of the twentieth century, anthropologists and cultural geographers attempted to define culture-areas. Diffusion studies sprang from this research. Culture-areas are defined as regions occupied by homogeneous culture groups. The first effort in this direction took place in 1895 when Otis Mason divided the Americas into eighteen culture-areas, based primarily on physical features (Harris 1968, 374). During the next thirty-five years, wide disagreement arose regarding the boundaries and definitional components of culture areas. Among the traits that anthropologists tried as bases for culture-area boundaries were food types, physical environment, developmental stage, and most popularly, language family.

Kroeber spent many years detailing the distributions of peoples
and artifacts in pursuit of accurately defining culture areas. While pursuing this goal, he makes vague references to diffusion paths, but no mention of diffusion mechanics are made (Kroeber 1920; Kroeber 1927; Kroeber 1939). In 1932, Kroeber and Driver attempted the ultimate in defining culture-areas, when they defined the culture-areas for Indian groups west of the Rocky Mountains in North America.

Having observed the presence or absence of almost 6000 cultural traits among the groups, they used their nominal data to generate coefficients of cultural similarity. Culture-areas were then defined in terms of these coefficients. No one has, as yet, attempted to repeat this process in any other region (Kroeber and Driver 1932; Harris 1968, 376-77).

Even with the use of highly detailed culture element lists, or perhaps because of the difficulty in obtaining them, anthropologists came to realize that, by either casually or meticulously placing boundaries, basic errors result in the process of defining boundaries. Four types of errors were seen as associated with the drawing of boundaries: "1) center and boundary change with [the] passage of time; (2) culture within the area may change so that it resembles cultures in different areas at different times; (3) portions of the area may be regarded as containing radically different cultures despite sharing of many features" (Steward 1955, 82); and (4) distant cultural groups within any defined area might be extremely different while the differences between adjacent, intermediary, groups are quite small. These small differences between proximate groups become substantial as distance and the number of intervening groups increase (Harris 1968, 377).
Clark Wissler in 1926 attempted to rectify these problems. He correctly realized that the problems were not with either the choice of boundary location or the criteria used to determine the boundaries, but with the culture-area concept itself. Wissler's scheme eliminated the culture-area and created the culture center. He defined a culture center as containing the "characteristic features of each area" (Harris 1968, 376). Wissler maintained that these characteristic features diffused outward from the culture center in every direction. The resulting landscape is one of cores and peripheries with inherent spatial variability in cultural elements because of incomplete diffusion rather than one of uniform regions where local variation is inexplicable. From these assertions, the age-area principle was not far behind (Kroeber 1931).

The age-area principle, first proposed by Wissler, states that the central point of a distribution of a cultural element is also the place where that element was invented. The age-area principle further states that the more widespread an innovation the greater its age, ignoring the possibility that rates of adoption can be affected by factors other than time and distance. The age-area principle was the first attempt to explain the attributes of the diffusion process (Harris 1968; Kroeber 1931).

2.2 Original Invention Versus Diffusion

From the attempts to define culture-areas, and particularly from Wissler's culture center concept, we see the beginnings of the principles used in diffusion research today. It took, however, another,
more controversial, research issue to establish diffusion studies and methods for the study of diffusion.

At about the same time that Wissler and Kroeber were trying to define culture-areas and culture centers, other anthropologists were trying to identify the original locations of major innovations (sedentism, agriculture, pastoralism). The anthropologists researching these innovations were divided into two camps. The first camp, the "diffusionists," claimed that the invention of an innovation occurred once and only once. The innovation subsequently diffused from that place and culture to the rest of the world and humanity. The other camp of anthropologists maintained that any particular invention could have been invented many times and diffused from many locations to nearby areas. These researchers cite historic examples in which different people invented the same thing at nearly the same time, such as Newton and Leibnitz, the inventors of the Calculus, or Darwin and Wallace, the discoverers of evolution. They believed that, when "the time was right" the inventing occurred where it was "needed" and that many people were capable of inventing an item or element. All that they required was just the proper circumstances and incentives.

The method employed by both camps to find the original locations of innovations, thus to prove their respective theories, consisted of cartographically implementing the age-area principle. Using historically or archaeologically derived dates of the first occurrences of innovations at many sites and then plotting both points and dates, researchers could draw isochrons, or lines of equal time. The most central isochron, the one encompassing the smallest area, necessarily contained the place of invention. If a single center was identified,
the diffusionists claimed that they were correct for that invention; if several distinct, isolated, isochrons existed, however, then those believing in independent innovation claimed they were correct (Childe 1937, 6). From the conceptual positions of both camps of anthropologists researching the origins of specific innovations, diffusion, per se, was still not a topic for research; rather diffusion was viewed as a condition that afflicted all places where the inventing had not occurred. Members of both groups agreed that the distributions of culture elements spread and were modified over space and time.

In the late 1940s, as anthropologists were beginning to lose interest, geographers began to undertake diffusion studies. This transfer began as early as 1928, when Carl Sauer collaborated with Alfred Kroeber on Pueblo Sites in Southeastern Arizona by Sauer and Brand (Platt 1952, 39). With the entry of geographers, emphasis began to shift from a simple search for the origin, to attempts to describe rates, direction, and spatial variation in the diffusion of innovations while continuing to use the cartographic method. One of the first works in this vein was "The Origin and Spread of the Grid Pattern Town" by Daniel Stanislawski in 1949. This article was followed by many others by various authors (Bruman 1948; Edmonson 1961; Gunter 1950; Kniffen 1951a; Kniffen 1951b; Isaac 1959; Jordan 1969; Kniffen 1965; Sauer 1952; Harlan and Zohary 1966).

The most notable example of prehistorical reconstruction in diffusion is Sauer's Agricultural Origins and Dispersals (1952). Sauer was also influenced by Kroeber in his diffusion studies of agricultural origins. Sauer's first article on agricultural origins was, in fact, one of a collection of papers presented to Kroeber on his
sixtieth birthday. A commonality in technique is obvious in works by these two scholars. Sauer, however, emphasizes the hearth and routes of diffusion much more than Kroeber.

More recently, Terry Jordan has used historical reconstruction to verify his claims regarding the origin of cattle ranching in Texas, in his article "The Origin of Anglo-American Cattle Ranching in Texas: A Documentation of Diffusion from the Lower South" (1969). In this article, Jordan describes the distribution and attributes of cattle ranching in the Southern United States and demonstrates that these same attributes are found in Texan cattle ranching. He also identifies the Texan cattle ranchers as migrants from the area of Southern cattle ranches. Jordan completes his argument by showing that coastal cattle ranching in Texas could not have come from the Spanish tradition because the Spanish did not herd cattle on the coastal plain (Jordan 1969, 81).

The titles of these works depict the continuing emphasis on locating the origins of innovations, or in Jordan's case, presenting an argument to support his chosen point of origin, and the developing secondary role of diffusion paths. The method used by the researchers in these articles were basic, historical and prehistorical reconstruction of a series of events. The events were then usually presented in a historic framework that was interpreted cartographically. Researchers emphasized the role of physical features, and recounted how the traits were diffused over, around, or through them. Historical descriptive articles following the spread of a particular innovation of this type had not long been produced before a method that sought general trends and regularity in the diffusion process was invented.
By the early 1960s, the origin-and-spread articles, which are still being produced, lost predominance in favor of research into the mechanics of diffusion. The new direction of research was the result of the quantification studies produced by Torsten Hagerstrand starting in 1953.

2.3 Hagerstrand and Spatial Diffusion

Torsten Hagerstrand was primarily responsible for the shift in emphasis that occurred in diffusion research among geographers. Hagerstrand laid the groundwork for this shift in his pioneer work on the adoption of farm implements, improvements, and production strategies by farmers in Sweden. Hagerstrand, unlike most of his predecessors, knew the origin of the innovations that he was studying and the paths of diffusion the innovations took. Armed with this knowledge, he hoped to explain how and why people at different places came to accept an innovation at different times.

Hagerstrand, while studying diffusion as a whole, emphasized the process of adoption, which he believed to be the crucial aspect of the matter. Adoption is the physical or mental acceptance of an innovation and, for technological innovations especially, is often identified by a purchase. Primary classification of individuals in the study by Hagerstrand reflects this emphasis. Hagerstrand, for the purposes of his diffusion analysis divided all individuals into two categories: adopters and potential adopters. Adopters are those individuals who have accepted an innovation. Potential adopters are defined as individuals who have not yet adopted the innovation, i.e. everyone else. According to the assumptions of the model, everyone
will eventually adopt the innovation.

By observing the spatial patterns of adopters and potential adopters at specified time intervals, Hagerstrand noted the growth of the adopter category at the expense of the potential adopter category. Observation of this phenomenon led Hagerstrand to state: "the spatial order in the adoption of innovations is very often so striking that it is tempting to try to create theoretical models which simulate the process and eventually make certain predictions achievable" (Hagerstrand 1967b, 7). Hagerstrand yielded to these temptations and constructed just such a model (Brown 1981, 18).

In his earliest attempts, Hagerstrand's model simulated the diffusion of the innovation evenly in all directions. He accomplished this by starting his model at the point and time immediately after the first adopter appeared from the midst of the ranks of the potential adopters. Hagerstrand used this person's location as the basis for the start of his simulation model. From the starting time the simulation progressed through a predetermined number of discrete time intervals of fixed length, representing some period of real time. In the simulation, this individual, during the next interval, contacted another individual and convinced him to adopt. During the next time interval each of the adopters contacted another, different, potential adopter and convinced them to adopt. The simulation continued until all potential adopters had become adopters.

Later, more accurate versions of Hagerstrand's simulation model gave detailed instructions on how an adopter contacts a potential adopter and converts him. Hagerstrand believed that the key to diffusion of an innovation was knowledge (Brown 1968). Potential adopt-
## Accumulated Intervals and Probability for Mean Information Field

<table>
<thead>
<tr>
<th>Interval</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-95</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>96-235</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>236-403</td>
<td>(0.0168)</td>
</tr>
<tr>
<td>404-543</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>544-639</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>640-779</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>780-1080</td>
<td>(0.0301)</td>
</tr>
<tr>
<td>1081-1627</td>
<td>(0.0547)</td>
</tr>
<tr>
<td>1628-1928</td>
<td>(0.0301)</td>
</tr>
<tr>
<td>1929-2068</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>2069-2236</td>
<td>(0.0168)</td>
</tr>
<tr>
<td>2237-2783</td>
<td>(0.0547)</td>
</tr>
<tr>
<td>2784-7214</td>
<td>(0.4432)</td>
</tr>
<tr>
<td>7215-7761</td>
<td>(0.0547)</td>
</tr>
<tr>
<td>7762-7929</td>
<td>(0.0168)</td>
</tr>
<tr>
<td>7930-8089</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>8070-8370</td>
<td>(0.0301)</td>
</tr>
<tr>
<td>8371-8917</td>
<td>(0.0547)</td>
</tr>
<tr>
<td>8918-9218</td>
<td>(0.0301)</td>
</tr>
<tr>
<td>9219-9358</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>9359-9495</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>9455-9594</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>9595-0762</td>
<td>(0.0168)</td>
</tr>
<tr>
<td>9763-9902</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>9903-9999</td>
<td>(0.0096)</td>
</tr>
</tbody>
</table>

**Cell Probability in Parentheses**
ters could only adopt after having been informed that such an innovation existed. When his conceptualization was implemented in the simulation model, Hagerstrand declared that every individual would be located in a cellular area and that each adopter could contact individuals within a two cell radius called a mean information field. Probability of contact within the mean information field was weighted such that an adopter was more likely to contact an individual in the cells nearest his own (Fig. 2.1). When the model was in operation, a scanner (a programmed device that is used to identify the location of adopters and upon identification to transfer control to the conversion routine) examined each cell once per simulated time interval. A simulated time interval might, therefore, conceptually represent a day, month, year, or century; in the actual operation of the model, however, each interval being simulated was represented as a single cycle of the scanner. If the scanner discovered an adopter in a cell, it centered a five cell by five cell information field over the adopter. This information field defined the adopter's limits of personal contact with others.

Each new simulation requires some means of determining which, otherwise identical, cell will hold the next adopter. A unique fixed range of values was assigned to each cell based on the distance of the potential adopter cell's centroid from the adopter cell's centroid. The decision was then made by selecting a random number and comparing it to the value ranges of the cells in the five-by-five mean information field. Sooner or later, the cell whose range contained the random number was found. Then the simulation assumed that, in that cell, a potential adopter was 'contacted' by the first adopter. The
model further assumes that the contacted potential adopter's cell became the site of an adopter and that he became a proselytizer on the next cycle of the scanner. The simulation continued until all potential adopters had been converted, or until the scanner had cycled some appropriate number of times.

Further improvement in the predictive accuracy of this initial, spatial model is achieved by constructing barriers that conceptually halt or slow diffusion. In a simulation a barrier is a wall of cells, or the juncture of adjacent cells, that deters the diffusion of an innovation. Barriers can be set up in several ways, but in practice to date they are rarely implemented systematically. Instead they are imposed in an ad hoc manner, and, except for barriers representing physical features, they have little or no theoretical basis or historical justification.

Several types of barriers have been defined. Barrier cells can accept the conversion signal and subsequently be skipped over when the scanner is searching for adopters to transmit during later time intervals, thus forming an absorbing barrier. Barrier cells can be formulated to redirect received signals to other cells lying only on the same side as the initial sender, thus creating a reflective barrier. A third type of barrier can be constructed by ordering a cluster of cells to wait one or more contacts before adopting, generating a 'conservative' zone where the innovation is slowed compared to surrounding cells. The first two types of barriers are examples of impermeable barriers, while the last is a permeable barrier. The attributes of any type of barrier can differ in space and time in the Hagerstrand model. For example, a barrier might be characterized as
being an impermeable, absolute barrier during the first ten intervals, as a reflecting barrier during the next ten intervals, and as no barrier during subsequent intervals (Yuill 1964).

Barriers are constructed to represent two different classes of phenomena. Barriers in the Hagerstrand spatial diffusion model can represent physical features (such as lakes, mountains, and deserts) that slow or halt diffusion because of the difficulty in communicating across them. The second class of phenomena is resistance to the innovation by potential adopters. Hagerstrand identified "a dichotomy of resistance as either the result of values which are inconsistent with adoption of the innovation (social resistance) or the result of practicalities which make adoption difficult or impossible (economic resistance)" (Brown 1981, 19).

Nearly all potential adopters have some degree of resistance to any innovation. "Few accept an innovation upon initial contact with it" (Hagerstrand 1967b, 263). Individual resistance, according to Hagerstrand, is a function of the sum of direct contacts made with adopters prior to acceptance by a potential adopter.

The results of Hagerstrand's model characteristically show one of two empirical regularities. These are the neighborhood effect and the hierarchical effect. The neighborhood effect appears most strongly when the study area is small and the unit of observation is the household. The neighborhood effect results in a strong tendency for the adopters to be clustered. Physical barriers act most disruptively when in the presence of a neighborhood effect. A cell adjacent to a strong physical barrier cannot readily establish contact with the comparable cell on the opposite side even if they both fall within
each other's mean information field. The neighborhood effect is specifically built into the Hagerstrand model because individual contact fields are limited to five-by-five cell areas (Hagerstrand 1967b, 159).

The hierarchical effect was not built into the original model. Hagerstrand first elucidated the principles of hierarchical diffusion in 1965 in his article "Aspects of the Spatial Structure of Social Communication and the Diffusion of Innovations," but it was Brian Berry who fully detailed the principles of hierarchical diffusion in 1971 with his explanation of the diffusion of television stations in the United States (1971). Berry discovered that hierarchical diffusion is most pronounced when dealing with diffusion over a large area among systems of cities. The hierarchical effect was first identified because of the inability of the original simulation models to predict the diffusion of innovations in such circumstances.

In hierarchical diffusion, the innovation is directly transmitted to the city with the highest innovation potential and then transmitted to other cities with successively lower innovation potentials. "The innovation potential of a center is a product of its position in the urban hierarchy and the force exerted on it by centers that have already adopted the innovation" (Berry 1971, 3). Hierarchical diffusion allows for the leap across space which occurs when distant cities adopt subsequently, but the intervening small communities and individuals do not. Conceptually, hierarchical diffusion occurs not at some linear distance, as it does when the neighborhood effect is in operation, but at an effective distance which is a measure of position in the hierarchy of places. Hierarchical diffusion models depict the
innovation as diffusing unidirectionally down the urban hierarchy, a level per time interval (Berry 1971; Hagerstrand 1965; Brown 1968; Crain 1966). It will be demonstrated, however, that hierarchical diffusion is not a special case of spatial diffusion, but a different kind of diffusion based in part on the degree of similarity between adopters. Further hierarchical diffusion, rather than arising from the influence of barriers, is a systematic modification of the distances between potential adopters. Strong theoretical support for the models implemented in subsequent chapters arise from this fact. Models that can help improve predictive accuracy and historical interpretation of diffusion simulation because they are based in these systematic modifications.

Hagerstrand was the first geographer to attempt to explain diffusion with the aid of quantitative models. His predecessors had largely limited themselves to qualitative descriptions of events which they observed during the diffusion of an innovation. Hagerstrand gained an advantage from his choice of innovations to study. Since he knew the history of the diffusion of the innovations that he studied, he was able to ask questions regarding how acceptance had occurred. Other researchers who did not know the history of their chosen innovation sought answers to the historical questions of where and when. Hagerstrand's choice of innovations permitted him to look outward from the origin of the innovation, in contrast to the historically oriented who sought to discover origins from distributions. When Hagerstrand looked outward from the origin, he saw an orderly progression of potential adopters becoming adopters with a spatial regularity approaching the spread of ripples on a pond, but the origin seekers,
who were focussed inward, conceived of the isochronous waves as pointing back to some point of origin.

2.4 The Network Diffusion Models of Anatol Rapoport

At the same time, that Hagerstrand was first publishing his model, in the early 1950s, Anatol Rapoport was studying network models. Rapoport, a biologist, initially used network models to describe the passage of information between neural cells in the brain. Rapoport depicted the brain as a graph with complex cycles. Each brain cell was described as a node and the ganglia the edges along which information is passed (Rapoport 1948a, 1948b, 1950a, 1950b, 1950c).

Rapoport claimed that the structure that he imagined within a brain also occurred within society (Rapoport 1953a, 1953b, 1954, 1956). This notion led Rapoport to construct a model remarkably like Hagerstrand's. The major differences lie in the implementation of the adoption process. Hagerstrand's model allowed an adopter to transmit an adoption signal only to a neighbor of the adopter. Rapoport's model placed each adopter and potential adopter in a cartesian coordinate system. Distances between individuals were measured and probability ranges assigned by the actual distances measured between individuals rather than by the fixed midcell to midcell distances that Hagerstrand used. The number of edges emanating from each individual varies based on the centrality of an individual in the group, the density of the group, the complexity of the group, and the perceived status of the individual. In denser and more complex networks, individuals are linked to a greater number of other individuals because the
number of associations between individuals increases with the density and complexity of the system. Those in leadership roles necessarily have more contacts than hermits. This fact is reflected in the number of edges emanating from each. The Rapoport model locates an adopter and randomly sends the adopt signal along an edge to a connected potential adopter.

Measured distances in the coordinate system between individuals need not be linear distances. They can be effective distances. For example, the x-y coordinate system might have available investment capital on the x-axis and number of hours spent reading financial news on the y-axis. Diffusion then would occur from an original adopter to potential adopters who had similar values on the x-y scales. In such a simulation, the variables portrayed on the axes become the only factors in the diffusion analysis. The ability to use non-spatial distances and to include more than two dimensions makes the Rapoport model a better choice for studying diffusion in a hierarchy than Hagerstrand's model.

Few have chosen to use Rapoport's models (Brown 1968). In the 1950s and 1960s, most geographers who were interested in diffusion were so enthralled by the Hagerstrand model that, had they known of Rapoport's models, they still may not have used them. When Brown discovered Rapoport's models in the late 1960s diffusion models were primarily being implemented by economic geographers who had progressed to supply side models. They were not very interested in consumer adoption. Cultural geographers, having explored the limitations of Hagerstrand's models, did not respond to Rapoport's contribution, perhaps assuming that it was covered ground. Had the Rapoport models
been implemented by cultural geographers on a large scale, the current emphasis on spatial diffusion would not be so great. The ability of network models to function on any x-y coordinate system and to use non-linear, non-distance coordinates would have revealed the explanatory limitations of spatial models.

2.5 Recent Trends in Diffusion Studies

After the brilliant work of Hagerstrand and Rapoport, quantitative diffusion research was neglected until the mid 1960s. Rapoport's work, not published in any social science journals, was easily missed by geographers until Brown brought his work to our attention in 1968. Hagerstrand, however, required translation before American geographers could appreciate his donation. In 1967 Allen Fred translated Innovationsforloppet ur Korologisk Synpunkt into English making it generally available to American geographers.

Before 1967, some American geographers used Hagerstrand's model. The University of Washington brought Hagerstrand to the United States to lecture for an academic year in 1959. During that year, Hagerstrand delivered his ideas to a group of quantitatively oriented students, several of whom went on to become premiere researchers in diffusion and the founders of the quantitative revolution (Brown 1981, 36).

During the 1960s, several articles were written on the mathematical principles and regularities of the Hagerstrand model. Yuill (1964), the first to perform this type of analysis, studied how barriers affect the shape of an innovation wave. He discovered that different types and shapes of barriers affect the spread of an innova-
tion just as they might affect any traveling wave phenomena. Yuill determined, therefore, that adoption of an innovation spread from a point outward in a wave.

Two articles by Morrill expound on the characterization of diffusion as a wave form passing through a population. "The Shape of Diffusion in Space and Time" (1970) discusses how the Hagerstrand model depicts the spread of an innovation using wave form parameters to describe the events at points at varying distances from the origin. In "Critical Parameters of Spatial Diffusion Processes" Morrill and Manninen (1975) show that an innovation can grow or die depending on the values of certain parameters, such as: number of originators, number of contacts a new adopter is permitted, and frictional factors. These articles, like those that follow up the work of any innovative thinker, serve to define and elucidate the principles first proposed by Hagerstrand.

The structure and utility of the mean information field was studied by several authors who agreed that it was a useful concept for the study of migration and information gathering processes (Morrill and Pitts 1967; Brown and Moore 1968; Mayfield 1967). Morrill and Pitts (1967, 422) make this point explicitly when they state, "Information fields, defined as measures of tendency to communicate over distance, have proven valuable as means of description and as tools in simulating the spread or movement of people or ideas."

Pitts spent much of the 1960s writing programs to implement the Hagerstrand models and developing programs and models based on Hagerstrand's model to perform diffusion simulation with non-uniform cells (Pitts 1965; Pitts 1967). Other geographers who have published simu-
lation diffusion programs are Zeller and Brown (1968), who published a
Markov-chain based model, and Marble and Bowlby (1968), who wrote
programs to analyze the results of diffusion models. These programs
have been little used because the few researchers who are capable of
running them are not inclined to use someone else's programs. These
published programs have, however, served to standardize and define the
rules and structures of Hagerstrand's models.

The majority of research on models of diffusion has been con­
ducted by economic geographers. These researchers have refined and
adjusted the models to suit their own developing theories. None has
been more active in this than Lawrence Brown who started in 1963 with
the first use of models based on Markov Chains. Most recently Brown
(1981) has developed the market and infrastructure context of diffu­
sion processes. According to Brown, Hagerstrand constructed his
spatial simulation model to depict the spread of an innovation among
consumers. Brown's adaptation facilitates modeling the diffusion of
the innovation among wholesale and retail merchants. Brown has also
identified processes relating to the initial phase of marketing by the
inventor. Brown's realization that consumers can generally adopt only
after merchants, has proved a fruitful concept for studying diffusion
in the modern world (Brown 1975; Brown 1981). Brown has done more to
advance the study of diffusion of economic and technological innova­
tions than any other researcher. He has achieved this by improving
the decision making features of simulation models to better reflect
corporate operations by identifying new applications for the models.
Simultaneously, however, he has promoted a focal shifting of diffusion
studies from cultural phenomena by developing techniques that are
Medical geographers, particularly those interested in contagious
diseases, have made great use of the models of Hagerstrand in the
past, but in recent years they have gone on to develop other models
which more accurately portray their subject (Haggett 1976; Girt 1978).
Other medical geographers have studied the diffusion of medical inno-
vations using spatial models (Coleman, Katz, and Menzel 1966; Baker
1979). Henry studied the diffusion of abortion clinics, using models
not unrelated to the models presented in chapter five, in the wake of
the 1973 Supreme Court decision permitting termination during the
first trimester (Henry 1978). Many studies performed by medical
geographers, however, are still performed using cartographic analysis
(Pyle 1972; Kwofie 1976). It is strange that some medical geographers
choose to employ only cartographic analysis because disease diffusion
truly lends itself to some form of spatial modeling.

Cultural geographers have used diffusion models infrequently and
have done little to improve them over the thirty years since Hager-
strand and Rapoport first constructed them.

Apparently, this is due largely to their having a differ-
et focus than researchers such as Hagerstrand. Often,
for example, diffusion studies were carried on with the
intent of establishing whether the configuration of the
cultural, social, or biological landscape is the result
of a diffusion process . . . for such an investigation the
actual process of diffusion or movement from one location
to another is not necessarily important (Brown 1968, 42).

Further reasons for the lack of use by cultural geographers lie
in the limitations of the models. The models of Hagerstrand and
Rapoport provide "simple mechanistic interpretations . . . for his-
torical reconstruction [which] can at best be only two dimensional"
(Hannemann 1975). How people come into contact with an innovation is
explained by the models, but why they accept is not. Economic geographers are not concerned with this issue, because it is obvious that a new device which is offered for sale is desirable. Cultural and historical geographers must constantly ask why the adoption occurred. What led people to accept an innovation? What are the effects of an innovation on the landscape? How is the innovation itself altered by acceptance by differing groups in differing locales? What other innovations are spawned by the adoption of the initial innovation? These are the questions highest in the cultural and historical geographer's mind. The need to ask why adoption occurred is particularly crucial to students of ideological innovations.

Geographers, on the whole, have accepted the spatial diffusion models without question. Many authors; especially those researching cultural, historical, and medical topics, have failed to use the models in their own diffusion research preferring, instead, to use cartographic analysis. The reasons are many and varied but it is known that cultural geographers were the last bastion of qualitative study in geography and that many were loathe to use mathematical models.

2.6 Summary

Diffusion research in American geography grew out of diffusion research in anthropology. Early geographic diffusion studies sought the origination point of the innovation and then discussed routes and spatial variability in innovation acceptance. Hagerstrand and Rapoport each constructed spatial diffusion models independently in the early 1950s (Brown 1968, 63). Each of their models allows for the
construction of barriers to diffusion. Hagerstrand's model is readily applicable to small groups of potential adopters in close proximity. Rapoport's model has advantages in the simulation of diffusion in a network of adopters who are linked to one another by researcher-defined criteria. Hagerstrand's model is limited to real distances except in the special case of hierarchical diffusion.

Since the initial development of the models, cultural geographers have used them sparingly, despite the great praise heaped upon the model as an example of scientific-analytical thinking in geography. Cultural geographers have done little to improve the models and their predictive abilities. The majority of research on the Hagerstrand model has been the search for and description of mathematical regularities in the results of the model. Of more import to economic geographers, the models have served to describe growth and spread of innovations to achieve a better understanding of modern economic structure.

The next chapter describes the construction, operation and forms of analysis for the specific model to explain the diffusion of prohibition from 1876 to 1919 in the spatial, socio-economic and combinatory dimensions.
CHAPTER THREE

A DIFFUSION MODEL FOR THE PROHIBITION MOVEMENT

Following the general discussion of spatial diffusion models and their historical development in chapter two, we may now describe the data and the diffusion model used for this analysis of the prohibition movement. The steps detailed below are taken in the production of the observational interrelation network, a simulation model, and the analysis of the results of the model for each set of x-y coordinates described in chapters four and five. Slight modifications to the routines permit the analysis of prohibition using x-y and z coordinates presented in chapter six.

This chapter begins with a description of the data used in the models, how the observational units are organized in reality and in the models, and what constituted the adoption of prohibition. The operation of the computer programs that perform the modeling and analysis are then described. Two programs are required to perform the diffusion simulations. The first program creates a network of edges linking counties based on selected criteria. The second program, described in the third section of this chapter, performs the actual simulation in the two dimensions specified. The last section of this chapter details the analytical techniques used to describe the accuracy and source of error in the predicted patterns of diffusion.

3.1 The Adoption of Prohibition

The first step in creating the model is the definition of the units of adoption, their organization and relationships, and the
characteristics of the adoption event.

Prohibition was adopted at every level of political organization between 1876 and 1919. In nearly all cases, the adoption of prohibition was achieved by political processes. Before 1919, the states held supreme authority for determining prohibition status, but most states empowered each of their counties to make the decision, via referenda. The New England states, New York, and several others scattered across the nation granted this power to units smaller than the county (township, town, or ward). Two states, Kentucky and Tennessee, kept this power to themselves and granted prohibition status to selected portions of the state when presented with petitions requesting a status change. Judges in Pennsylvania circumvented the political process and declared counties dry by their judicial authority. Each state reserved to itself the power to change its status either via statute or constitutional amendment.

In most states, adoption of prohibition at the state level occurred only after a significant proportion of the counties in that state had adopted. State level adoption in these cases represents a hierarchical leap. A hierarchical leap occurs when lower order units in the hierarchy of places, people, or things influence a higher order unit. Not only was adoption by each of the states a hierarchical leap, but national adoption represents another hierarchical leap. One in which the adoption by the states influenced the actions of the federal government.

Like all innovations, the diffusion of prohibition occurred primarily as a result of the actions of change agents (proselytizers). The upward hierarchical diffusion was the result of the actions of
these vocal campaigners. These people first campaigned to dry up their community, then their county, state, and nation. For some, success at the local level was enough, and they dropped out of active support for the movement, but for others, once they were involved, success meant a broadening of horizons and new territory to be conquered. The conquest of liquor was necessarily up the hierarchy because most campaigners were locked into a particular unit at each level in the hierarchy.

The observed hierarchical organization of the political units is vastly simplified in the model. Only two units, the county and the nation, are represented. The other units are reclassified using two rules. First, in states where the adopting units were smaller than the county, all the units within a county must have adopted individually before the entire county is considered dry for the purposes of analysis. This is a stringent, but necessary, criterion designed to maintain consistency in the data across the nation. The majority of sub-county units having adopted is not a viable criterion for placing the entire county in the dry category because of the strong tendency for the rural units to adopt first. In such a case, the majority of the county population may be located in wet territory even though the majority of townships and land area in a county are dry. Second, when an entire state was dry, each of its counties was an adopter, just as when the nation went dry all states and counties were compelled to adopt.

Given these rules, data regarding the prohibition status of each county were gathered on an annual basis for the period 1876 to 1919. During data collection, each county was assumed wet (potential adopter
status) until contrary evidence was provided. The primary data sources for the period 1876 to 1903 are historical accounts of the movement in each state. These histories focus on the activities of the proselytizers and their adversaries, not on the extent or location of dry territory. Nearly all the histories include at least one map depicting dry counties within their geographic frame for a selected year, and give lists of dry counties for selected years. From these maps and lists the status of the depicted counties was ascertained for that year alone; no attempt was made to extrapolate. Much information regarding which counties were dry was gleaned from these histories. Data accuracy is vastly improved after 1904. When the Anti-Saloon League began publishing a yearbook that included both maps of dry territory by county, for the nation as a whole, and by township for selected states. The yearbooks also included accounts of territorial gain during the preceding year. The data collected are encoded in binary format, indicating if a county was wet or dry, and stored in the computer in anticipation of the analysis to follow.

The nation is defined as the set of coterminous counties. The national hierarchical leap is built into the model and occurs when over eighty percent of the counties are predicted dry. Eighty percent was chosen as the threshold value because in 1918, on the eve of national and complete adoption, 80.627 percent of all counties were observed dry. All models are contrived to stop their iterations once the model's predicted percent dry surpasses this threshold. The advantage in this arbitrary decision lies, in addition to its having at least some warrant in history, in the savings achieved in computer time. The computer time to calculate adoption in the last twenty
percent is substantial because the rate of adoption slows as the system becomes saturated. Not allowing the machine to calculate the final portion of the adoption sequence has no ill effect in the following analysis.

3.2 Preparation of Input Data Dimensions

The x-y coordinates of the counties are used as the basis for creating a network (multiply linked list) depicting the relationship of the counties in the selected data dimensions. X-y values are derivable from any number of variables. Program 1 accomplishes this task. The details of how the network is created, the input and output file structures, and the program parameters are described here. An understanding of how the input file for the simulation model is created and organized is imperative for comprehending the simulation model.

Program 1 (Appendix) generates a list of the nearest twelve counties (Fig. 3.1) for each county in its input data file. The choice of twelve potential adopters deviates from the classic Hagerstrand MIF, which was composed of 25 Swedish Survey System grid cells. Rather than impose a grid cell system upon the counties of the United States, forcing several counties into a single cell while leaving other cells empty, a network information structure was selected. The network structure is appropriate for all Cartesian coordinate systems where each point represents an observation, while the Hagerstrand type MIF is designed for simulating diffusion with multiple potential adopters per cell.

Program 1 requires four input fields: 1) an identification code
**Figure 3.1: Flowchart of Program 1, Data Structure Creation**

```
read in all observations i

i=1

i=1+lastobs

yes

output obs i

vars ICPSR,
SECTOR1-12,
PROB1-12, Y876

i=i+1

no

j=1

j=1+lastobs

yes

sum= $\frac{i}{n+1} \text{PROB(n)}^2$

n=2

no

PROB(n)=PROB(n)+PROB(n-1)

n=n+1

no

j closer

** Is obs j near obs i than
any previously encountered
obs j

no

calculate distance
between obs i & j.
calculate angular position
of obs j relative to obs i
and assign to a sector

j closer

RECORD STRUCTURE (INPUT)
ICPSR, x-coordinate, y-coordinate, Prohibition Status 1876 (Y876), IDNO

RECORD STRUCTURE (OUTPUT)
ICPSR, ptr Sector1 (SECTOR1), Probability Sector1 (PROB1), ptr Sector2 (SECTOR2), Probability Sector2 (PROB2), ptr Sector3 (SECTOR3), probability Sector3 (PROB3), ....Ptr Sector12 (SECTOR12), Probability Sector12 (PROB12), Prohibition Status 1876 (Y876), IDNO

(Record numbers are array positions of counties in the output data file when written out here and when read in to Program2.)
```
for each county, ICPSR\(^1\); 2) the prohibition status of each county in the year 1876, Y876; 3) the county's x-coordinate, and 4) the county's y-coordinate. The x and y coordinates are given different names, depending on the data dimensions that they represent. The prohibition status of counties in 1876 is constant across all runs of the simulation program. For the analysis, however, the status of each county must be known for each year in the study period. The output of program1 has 27 data fields per record. The ICPSR code and Y876 are stored by program1 and printed out unaltered. A sequence number, IDNO, is given to each county in the list of records. Of the remaining 24 data fields, SECTOR1 through SECTOR12 hold IDNO for the closest county in each of the twelve thirty degree sectors that surround each county. The other twelve fields, PROB1 through PROB12, hold the probability of contact for each of the twelve object counties, identified in the SECTORn fields, during each simulated period.

The algorithm for program1 is simple in concept. To put the algorithm into use, however, takes a great deal of time because every county's coordinates must be examined in relation to every other county's coordinates. The county for which neighbors are being sought is called the subject county, and all other counties are called object counties. Distance between subject and object counties is calculated using the Pythagorean Theorm. For each subject county, a quadrilateral centered on that county is searched for other counties. The quadrilateral is of some predetermined size that varies with the dimensions used. For variables that encompass wide ranges, it is

\(^1\)ICPSR stands for the Inter-university Consortium for Political and Social Research. This study uses the ICPSR county identification code.
necessary to adjust the search parameter so a larger absolute area is searched when attempting to locate object counties. If this ad hoc adjustment is not made, many sectors will point to the null county and the network will be disjointed. Alternately, if the search area is too large, not enough counties will be culled by this subsetting criterion, and counties will have to be rejected by other, more computationally expensive, criteria.

Counties that fall within another's quadrilateral search area are then tested for inclusion in a circle whose radius is half the length of a side of the quadrilateral. In the spatial dimensions, the search area is not circular but elliptical because as latitude increases the length of a degree of longitude decreases; the the total deviance, however, when taken from the mean latitude is negligible. Each county seat discovered within the circle necessarily falls within one of twelve thirty degree sectors. Once an object county is identified as being within the circular search area, its angular direction relative to the subject county is calculated and classified into one of the twelve sectors. If that county is closer to the subject county than any other object county previously encountered in that sector, its IDNO and distance are stored in the SECTORn and PROBn fields for that subject county. In effect, an edge is drawn between the SECTORn object county and the current subject county. Each edge is a potential path for information to travel along.

If, for example, the search program is attempting to identify the nearest neighbors in the spatial dimension's coordinate data set for Chester County, Pennsylvania, IDNO 2189, it first identifies all county seats within a one degree quadrilateral of West Chester, the
seat of Chester County (Fig. 3.2). In the spatial dimensions, the latitude and longitude of the county seats are used as the x and y coordinates for input to the diffusion model. The search area for the quad in the spatial dimensions is set at degree of latitude and longitude on a side. The radius for the search circle is, therefore, a degree. Any county seat within the square, but not the circle is rejected, as is the case with Somerset County, New Jersey. Counties are examined in IDNO sequence by the rejecting criteria.

The first county encountered that is not rejected is Kent County, Delaware, IDNO 286. Calculations determine that Kent County is in SECTOR6, and since no other SECTOR6 counties have been previously identified, its IDNO is placed in Chester County's SECTOR6 field, and the distance between their county seats is placed in Chester County's PROB6 field. The next county that meets the criteria is New Castle County, Delaware, IDNO 287. New Castle is also located in SECTOR6 and, since it is closer to West Chester, its information replaces Kent County's information in the SECTOR6 and PROB6 fields of Chester County. The third county that passes the tests is Cecil County, Maryland, IDNO 1142, which is not only the closest SECTOR8 county found so far, but it is the closest overall. After Hartford County, Maryland, IDNO 1147, is assigned to SECTOR9, we see that Kent County, Maryland, IDNO 1149, even though it is within the square, is further from West Chester than the seat of Cecil County. Kent County, Maryland is, therefore, not added to the list of counties proximal to Chester County. Sequencing through the counties by IDNO continues until all counties have been processed and the closest county seat to West Chester in each sector has been identified (Table 3.1).
Spatial Distribution of County Seats About Chester County, Pa.
Table 3.1: Typical Record Generated by Program1 for Spatial Dimensions

Record 2189: Chester County, Pennsylvania

<table>
<thead>
<tr>
<th>Sector</th>
<th>Pointer1 Value</th>
<th>Pointer County</th>
<th>Distance</th>
<th>1-dist.</th>
<th>Each</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 NNE</td>
<td>2231</td>
<td>Lehigh, Pa.</td>
<td>.6379</td>
<td>.5931</td>
<td>.0810</td>
<td>.0810</td>
</tr>
<tr>
<td>2 NE</td>
<td>2220</td>
<td>Montgomery, Pa.</td>
<td>.2968</td>
<td>.9119</td>
<td>.1246</td>
<td>.2056</td>
</tr>
<tr>
<td>3 ENE</td>
<td>2225</td>
<td>Philadelphia, Pa.</td>
<td>.3932</td>
<td>.8454</td>
<td>.1155</td>
<td>.3211</td>
</tr>
<tr>
<td>4 ESE</td>
<td>2197</td>
<td>Delaware, Pa.</td>
<td>.2040</td>
<td>.9584</td>
<td>.1310</td>
<td>.4521</td>
</tr>
<tr>
<td>5 SE</td>
<td>1714</td>
<td>Cumberland, N.J.</td>
<td>.7433</td>
<td>.4475</td>
<td>.0612</td>
<td>.5133</td>
</tr>
<tr>
<td>6 SSE</td>
<td>287</td>
<td>New Castle, Del.</td>
<td>.2062</td>
<td>.9575</td>
<td>.1309</td>
<td>.6442</td>
</tr>
<tr>
<td>7 SSW</td>
<td>000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.6442</td>
</tr>
<tr>
<td>8 SW</td>
<td>1142</td>
<td>Cecil, Md.</td>
<td>.4357</td>
<td>.8102</td>
<td>.1107</td>
<td>.7545</td>
</tr>
<tr>
<td>9 WSW</td>
<td>1147</td>
<td>Hartford, Md.</td>
<td>.8117</td>
<td>.3411</td>
<td>.0466</td>
<td>.8015</td>
</tr>
<tr>
<td>10 WNW</td>
<td>2210</td>
<td>Lancaster, Pa.</td>
<td>.5356</td>
<td>.7513</td>
<td>.1027</td>
<td>.9042</td>
</tr>
<tr>
<td>11 NW</td>
<td>2180</td>
<td>Berks, Pa.</td>
<td>.6741</td>
<td>.5456</td>
<td>.0746</td>
<td>.9788</td>
</tr>
<tr>
<td>12 NNW</td>
<td>2187</td>
<td>Carbon, Pa.</td>
<td>.9192</td>
<td>.1551</td>
<td>.0212</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

1 This column of data is stored in the field SECTORn
2 This column of data is stored in the field PROBn
Program1 next sums the square of the values held in the twelve PROB fields. Deterioration rates in phenomena affected by distance decay, generally occur at one over distance squared, the intensity of light, for example. One over distance squared was adopted as the rule for the model because information flow is a distance decay phenomenon, and it is assumed to behave like other distance decay phenomena. The probability of contact from the central point to any of the twelve points was derived for each of the twelve points using the formula

\[ P_n = \frac{(1-D_n^2)}{\text{SUM}}. \]  

(1)

\( P_n \) is the probability of the county in \text{SECTOR}n being contacted in any generation. \( D_n^2 \) is the squared distance between the central county seat and the county seat, connected by an edge, in \text{SECTOR}n. \( \text{SUM} \) is the summed squared lengths of all edges from the central point.

The probability values for the county in each sector are then added to obtain a cumulative probability value for each. For example, the probability of contacting the county held in Chester County's \text{SECTOR1} field, Lehigh County, Pennsylvania, is .081 and is stored in Chester County's PROB1 field over the distance that was previously stored there. The range of values represented in the PROB1 field are .000 to .081. The probability of contacting the county held in \text{SECTOR2}, Montgomery County, Pennsylvania, is .1246, but the value stored in PROB2 is .2056 (.081 + .1246). The range of values represented in the PROB2 field are .081 to .2056. The probability of contacting Philadelphia County, Pennsylvania, \text{SECTOR3}, is .1155, and the value stored in PROB3 is .3211 (.081 + .1246 + .1155). The PROB3 field therefore represents the values between .2057 and .3211. (The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
value stored in PROB12 is always in excess of .99, but less than 1.0, owing to rounding error.)

The storing of the cumulative probability value in the PROBn fields facilitates the search for the sector that is selected for adoption in the simulation. To simulate the diffusion of an adoption a random choice of direction must be made. This is done by generating a random number between .0001 and 1.0 from a uniform distribution. This random number is then compared successively to each of the values in the PROBn fields. The first PROBn field whose stored value is greater than the random number identifies the county stored in SECTORn as having been selected for adoption. The probability range of this PROBn field contains the random number. This selection process insures a random choice of direction for growth at each juncture, while it permits, after sufficient iteration, the innovation to spread to all counties in the simulation.

3.3 Nature and Organization of the Diffusion Model

In principle, this diffusion model and the program that implements it (Appendix), are identical to Hagerstrand's diffusion model and Rapoport's network model (Fig. 3.3). The differences in the structure of the MIF do not substantially alter either the form or the concept of the model. The major differences that exist between the model used here and the model used by Hagerstrand and others are the characteristics of the adopting units. Hagerstrand's model assumes that multiple adopters live in each areal unit.

Here the areal units, individual counties, are occupied by but one organism—the duely constituted government of the county itself.
Figure 3.3: Flowchart of Program 2, the Simulation Model

initialize random number generator with seed value

read output of program 1

totdry = 0

todry < .8

do for all obs.
z = current obs.
country z dry

draw random # x = 1

do while random # > probability value of sector(x) x = x + 1 (x is always < 13)

when sector with prob value > random # county residing in z's sector(x) ptr field is set to dry status

output iepsr and result

totdry = totdry / N

totdry = totdry + 1

do for all obs. i = current

concatenate results from most recent loop to previous results for county i

RECORD ORGANIZATION (INPUT)
ICPSR, P tr Sector 1 (SECTOR1), Probability Sector 1 (PROB1), P tr Sector 2 (SECTOR2), Probability Sector 2 (PROB2), P tr Sector 3 (SECTOR3), Probability Sector 3 (PROB3), ..., P tr Sector 12 (SECTOR12), Probability Sector 12 (PROB12), Prohibition Status 1876 (y876), IDNO

RECORD ORGANIZATION (OUTPUT)
ICPSR, concatenated string of predictions (RESULT)

(Record numbers are array positions of the county when read in. Records are stored in record number order in the data file.)
The county, therefore, represents both the sole adopter and the unit of adoption. It is represented by a point as in Rapoport's model. When simulating diffusion in the spatial dimensions, the model uses points corresponding to the latitude and longitude of each county seat.

Turning to more realistic models, social, economic, or other qualities of places can be used as the 'dimensions' that provide values that determine the organization of the network of observations. The point used to represent a county may be that county's value on two social, economic, or other type of indicators. Examples of valid variables for use in the model include: county population, county population density, median individual income by county, percent of population Roman Catholic, or an index consisting of a combination of these. The potential for using any two ratio scale variables in this model makes it much more flexible and useful than a model that can only use spatial references as positional parameters for generating the connectivity network.

Because this model is nearly identical to the Hagerstrand model, its results have the same properties. These properties have certain advantages and disadvantages when the results of using the model are compared to historic fact. One of the model's properties that affects the results is the rate of adoption. The rate of innovation adoption is ordinarily initially slow, rapid in the middle of the period, and slow again when the population of potential adopters have nearly all adopted. Once near saturation is achieved, growth in new adopters becomes nil. In the case of prohibition, the saturation slowdown did not have a chance to occur because the Federal government, in a hier-
archial leap, adopted national prohibition in 1919.

A second property of the model is that it can predict neither periods of no diffusion nor reversals to disadoption. During the course of the actual events, there were periods of recession or retreat from dry status. The model, however, can only predict an increase in the number and area of adopters. Years that the model makes predictions for, therefore, occur at odd intervals because years when there is no net gain in adopters cannot be identified and are treated as nonexistent by the model.

Third, once a county has been predicted dry by the model, it cannot be subsequently predicted wet. This rigidity results in a large number of errors because many of the counties that were dry in 1876 reverted during the middle years of the study period. Other counties oscillated between wet and dry on an annual basis. It was decided to maintain the dry status in these counties and accept the errors when analyzing the model's results, rather than cancelling them out of the analysis.

Program2, the simulation program (Appendix), begins by reading in, for each county, the county sequence number, IDNO; the county ICPSR code, ICPSR; pointer fields, SECTOR (SECTOR1 - SECTOR12); probability fields, PROB (PROB1 - PROB12); and prohibition status in 1876, Y876. A unique seed value, which in turn causes the generation of a unique sequence of random numbers, is specified for each run of the program. After reading in the above fields, program2 creates three additional fields for each county: TEMP, NEX, and RESULT. Initially the value of Y876 is copied into each of these three fields. TEMP and NEX are single character alpha fields. RESULT is a forty character
alpha field.

After all data fields are initialized, program2 enters into its simulation sequence. In the simulation, a scanner is activated which examines each county's field TEMP for the condition TEMP='1', which indicates a dry county. When the scanner encounters a dry county, it temporarily passes control of the program to the conversion section of the program. As long as the scanner finds wet counties, TEMP='0', it continues to scan down the IDNO sequence until it has examined the entire sequence.

The conversion section of program2 decides which of the counties adjacent to the dry subject county will become dry. The conversion process is accomplished by first drawing a random number, which has been generated by the random number generator. The random number generated is from a uniform distribution and ranges between 0.0001 and 0.99. The basic algorithm (Appendix) for computing the random number sequence is

\[ \text{RAND}_{n+1} = \frac{(\text{RAND}_n \times 899)}{32767}. \quad (2) \]

The initial value of RAND\(_n\) is called the seed value, and all subsequently generated random numbers are a function of its value. The probability fields of the dry county are searched sequentially (PROB1, PROB2, PROB3, ..., PROB12) for a value greater than that of the random number. After the search has revealed this PROBn field, the search is halted, and the field NEX of the county corresponding to the contents of the SECTORn field is given the value '1'.

From the previous example let us assume that, county 2189 is identified as being a dry county by the scanner, and the random number
drawn is 0.7132. In county IDNO=2189's PROB fields are the following values: PROB1=0.0810, PROB2=0.2056, PROB3=0.3211, PROB4=0.4521, PROB5=0.5133, PROB6=0.6442, PROB7=0.6442, PROB8=0.7545, PROB9=0.8015, PROB10=0.9042, PROB11=0.9788, PROB12=0.9997. PROB1 through PROB7 are each examined, and they leave the condition unsatisfied because the contents of each is less than 0.7132. PROB8 meets the condition because the value it contains is greater than the generated random number. The value held in SECTOR8 is 1142. The field NEX of the county, whose IDNO is 1142, is given the value '1'.

When the scanner has completed its examination of all counties, it passes control to the update section of the program. In the update section, the value of NEX for each observation, IDNO, replaces the value that is held in TEMP, and the value of NEX is also concatenated to the value of RESULT. After all observations have been updated, control is passed back to the scanner.

The method of storing changes made during a scan cycle in NEX, while searching for dry status in the field TEMP, eliminates the possibility of entering the conversion section for a county that has become dry during the same scan cycle. If there were only the field TEMP, then dry county, IDNO=2189, would, in the previous example, have caused the value of TEMP for county, IDNO=1142, to change. When the scanner subsequently examines TEMP of IDNO=1142, it would find a dry status indicator and would pass processing control to the conversion section. As the scanner proceeds down the sequence of IDNOs, the number of premature adoptions would expand exponentially.

Once eighty percent of all counties have been declared dry, as simulated by the model, further scanning and updating stop, and con-
trol is passed to the output section. The identification of the percentage of counties predicted dry is a subfunction of the update section of program2. The output section simply prints, for each county, the ICPSR and RESULT fields. The program then terminates.

3.4 Methods Employed in Analyzing and Interpreting the Results of the Diffusion Model

After the model has been run, it is necessary to verify the model results. Verification is achieved by comparing the model results to the events observed in the real world. This technique permits us to determine how well the model mimics reality. The verification process also identifies the errors made by the model. The type and position of model errors identifies not only the failings of the model, but they reveal unexpected and previously unnoticed trends in the growth of prohibition. The operations described here are crucial in the acquisition of new information regarding both prohibition and the diffusion model.

Regardless of the data dimensions used in the generation of the model, the following steps are taken to analyze the results after the fifth run of the simulation model with five different seed values for the random number generator. The model must be run several times because, like the Hagerstrand model, it is stochastic in nature. Stochastic means that the results vary with the random numbers generated and used by the model. Each sequence of random numbers produces idiosyncratic results because different sequences result in potentially different choices in the sectors, and therefore in object counties. These idiosyncrasies must be removed before the true predictive
ability of the model can be determined. Their removal is accomplished by averaging the results of several runs. The mean result of several program2 runs is called the omega pattern. The omega patterns in all models run in the analysis here are created from five individual runs.

The first step of the analysis is, therefore, the generation of the omega pattern. The merging routine (Appendix, program3) reads the fields ICPSR and RESULT from each of the five data sets for each county. Program3 then adds the contents of each column from all five of the RESULT fields and stores the summation in a sixth forty column result field, OMEG, each column's value ranges from zero to five. A zero in column one of an OMEG field indicates that none of the runs predicted that county dry in the first scanner pass. A five indicates that all five runs predicted the county dry during the first scanner pass. For example, for county IDNO=1142 the first ten columns in the field RESULT for the first run equalled '0000111111', for the second run '0001111111', for the third run '0000001111', for the fourth run '0000011111', and for the fifth run '0001111111'. OMEG for the first ten time intervals for that county equals '0002345555'. In subsequent analyses, columns with '3' or greater in OMEG are considered to represent dry periods, and columns that equal '0', '1', or '2' represent wet periods (Fig. 3.4).

After the omega pattern has been generated for a set of data dimensions, it must be matched to the observed pattern. This task is accomplished by comparing the percentage of counties predicted dry in each column of the omega pattern to the percentage of counties observed dry for each year of the study period. Because the number of observations differ from one set of data dimensions to the next, it is
Figure 3.4: Flowchart of Program3, Merging-Omega File Creation

do for all observations

read obs i from all 5 files containing output from the 5 simulation runs of a model

j=1

omega(j)=sum of all five predictions found in column j for observation i from each of the five files

j=j+1

j=40

yes

no

end of input files

yes

end

RECORD ORGANIZATION (INPUT, all five files)

ICPSR, output of a simulation run using a unique random number generator seed value (RESULT)

RECORD ORGANIZATION (OUTPUT)

ICPSR, merged result of all five simulation runs using different random number generator seed values (OMEG)
necessary to derive the percentage predicted dry in each column of the
omega pattern and to recalculate the percent observed dry for each set
of dimensions. Each omega pattern column is then assigned to the
observed year whose percentage dry most nearly equals the predicted
percentage dry term of the omega pattern. The correlation between
observed status and predicted matches is calculated to identify the
degree of fit between them. The correlation is used to verify that
each omega pattern matches reality equally, or nearly equally, as well
as the others. Equality, or near equality, is important for compari-
sions between omega patterns and for limiting error in the temporal
dimension within an omega pattern. When the correlation between the
observed and predicted temporal patterns is very high (.99) then one
source of error has been virtually eliminated. Distributional error
terms are all that remain for identification and explanation.

Each successive column of predictions in the omega pattern fore-
casts a greater percentage dry than its predecessor. To obtain a
match in percentages, therefore, predictions can necessarily only
match observed years whose percentage of counties dry is greater than
any previously observed percentage. Matches between omega pattern
columns and observed years occur at irregular intervals, mostly
because years showing growth in the percentage of dry counties were
interspersed with years of no growth and gains in wet territory.

The rate of growth can outstrip the rate of predicted expansion
in the number of dry counties. This may occur because the effects of
saturation are being felt in the model more than they are in reality.
When this situation arises the predictions that match best are used
and the intermediary predictive generations are dropped from further
consideration. For example, suppose omega pattern column eleven predicts 1873 dry counties, column twelve predicts 1947, column 13 2009, column 14 2065, and column 15 predicts 2100 dry counties. In 1914 there were 1802 counties observed dry, in 1915 there were 1896, and in 1916 there were 2105. The matching algorithm would match column eleven with 1915 and column 15 with 1916. Omega pattern columns 12, 13, and 14 are interpreted as representing the status of prohibition territory at three distinct points in time during the year 1915. Because observed data are not available for portions of the year, however, these columns are dropped from the remainder of the analysis just as years with no matching columns in the omega pattern are dropped.

For each county in each omega year, the analysis program compares the predicted status to the observed status. Four possible codes exist to describe the result of the comparison. The code '0' is used to describe a county that was wet and was predicted wet. The code '1' describes counties that are dry and predicted dry. Type I errors\(^2\), counties that are wet but predicted dry, are coded '2'. The final category, type II errors, consists of counties that are observed dry and predicted wet. These counties are given the code '3'. The codes '0' and '1' represent correct predictions, and the codes '2' and '3' represent model errors for those counties in the omega year where they occur.

\(^2\)In statistical terminology, a type I error occurs when the true state of nature in a binary condition is misidentified in a specified direction. A type II error occurs in the same situation except that the opposite state of nature is true and misidentification has also occurred.
For each omega year the percentage of counties predicted correctly is calculated. The accuracy of the first year predicted is always highest because this year is most like 1876, and the model has not yet had the opportunity to make many mistakes. Measured accuracy decreases until the system approaches saturation. With approaching saturation, opportunity for error decreases because of the high level of agreement in the number of counties predicted dry. If the accuracy readings are graphed over time, a parabolic pattern will always emerge. The overall accuracy of each model is obtained by finding the mean of the individual year's accuracies. This single figure summarizes the predictive ability of each set of variables.

The greater the predictive ability of a set of variables, the more the diffusionary pattern derived from them, via the model, corresponds to the observed pattern of diffusion. The network derived from a set of variables defines the interrelationships of counties in that coordinate system. There exists, it is assumed, a set of variables from which a coordinate system can be constructed that will exactly match the observed pattern of growth. The identification of this set of variables must be the ultimate goal of this research, even if for the time being, the size of the set is limited to a very few variables.

Examination of the distributions of each code value, 0, 1, 2 or 3, over time provides information regarding both systematic and random failings of the models. The primary tools for determining the distribution of each code value are summary tables and scattergrams. When a series of related errors in prediction is identified, an explanation is sought and offered. Non-random predictive errors are associated
with an event that has modified either the rate or direction of growth from the expected as defined in the model. The identification of these events tells much regarding the sensitivities of the movement.

Random errors, on the other hand, are linked to the model's method of choosing counties for conversion. During each pass of the scanning algorithm, each dry subject county transmits a signal to one of its object counties. When this signal is sent to a county that does not, in reality, go dry that year an error has been made. This error is only considered random, however, if another object county of the same subject county does indeed go dry that year. Otherwise the error is classified as systematic in nature. The error is random because it was right to select one of the object counties of that subject county for adoption, but the random number drawn did not select the proper object county. Each random error, by definition, therefore, creates one type I error (the wet object county incorrectly predicted dry) and one type II error (the dry object county that is still predicted wet).

Each omega pattern is additionally analyzed for accuracy by region using the above four codes. These regions are the New England states, the Mid-Atlantic states, the old Northwest Territory states, the Plains States, the Mountain West states, the Pacific West Coast states, the North-South Border states, and the Southern states. New England consists of Maine, Vermont, New Hampshire, Connecticut, Massachusetts, and Rhode Island. The Mid-Atlantic states are New York, Pennsylvania, New Jersey, and Delaware. The Northwest Territory is Ohio, Indiana, Illinois, Michigan, and Wisconsin. The Plains states are Iowa, Minnesota, North and South Dakota, Nebraska, and Kansas.
The Mountain West states are Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, and Nevada. Washington, Oregon, and California are the Pacific West Coast states. The North-South Border states are Oklahoma, Missouri, Kentucky, Tennessee, West Virginia, and Maryland. The South contains more states and counties than any other region. This region's member states are Virginia, North and South Carolina, Georgia, Florida, Alabama, Mississippi, Louisiana, Arkansas, and Texas.

The regional analysis of the distribution of code values aids in the identification of spatial trends in prediction because subsetting the population regionally permits clusters of errors, and correct predictions, to emerge distinctly. The number of dry counties in a region, for example, may be substantially different from the predicted number because the number of predicted dry counties is matched to the national number observed dry and not the regional values. The level of over and under prediction for all regions offsets each other when agglomerated into the nation. This is as true for non-spatial variables as it is for spatial ones.

3.5 Summary

The data and methods for producing and analyzing the results of a stochastic model that simulates the diffusion of prohibition in terms of any two variables in the United States from 1876 to 1919 have been described. Before the diffusion model can be put to work on a set of data, however, it is necessary to calculate the distances, in the appropriate dimensions, between observations. From the point-to-point distances, the nearest neighbor of each county is identified in each
of twelve thirty-degree sectors. The distances are converted into probabilities of interaction during each current interval in the simulation, based on the distance squared between the subject county and each of the twelve nearest object counties. When for each county, its nearest neighbor in each of the twelve sectors has been identified and the probability of it contacting each calculated, the simulation program can be run.

The simulation program models the diffusion of prohibition by locating dry counties and having them proselytize among the counties identified as being closest in each of the twelve sectors; only one neighbor can be converted during each simulated interval by each dry county. The counties that were actually dry in 1876 are so identified. These counties are the seeds from whence the movement grows in subsequent years. After each simulated interval, more counties have been converted from wet to dry status. The program eventually transforms eighty percent of all wet counties to the dry status before terminating.

The predictions made by the model are first analyzed by matching the predicted and the observed diffusion status of each county. The percentage dry predicted in each interval is identified and matched to the observed year with the closest percentage observed dry. Accuracy is then defined as the number of observations whose predicted status in that year is the same as the observed status divided by all observations. The overall ability of the simulation to predict status is found by averaging the percent accuracy for prediction years. The organization of the model states that the object counties that are closest to a dry subject county, in the appropriate dimensions, have
the highest probability of adopting next. In many cases, therefore, the shortest link, or edge, is the predicted path of diffusion because the selected random number is most likely to fall within the range associated with that object county. The accuracy calculation and presentation of findings are followed by a distributional analysis of the errors. The error analysis, both nationally and regionally, tells when and where the model does not conform to reality. At these times and places the model operates using decision making criteria that do not reflect reality. Once these times and places are identified explanations for failure to predict observed adoption or non-adoption can be offered beyond that supplied in the model.
CHAPTER FOUR

ANALYSIS OF SIMULATED DIFFUSION IN THE SPATIAL DIMENSIONS

In this chapter, the analysis begins. The three previous chapters have readied the reader to evaluate critically what the model does and does not reveal about diffusion processes and the prohibition movement. Of the competing sets of variables, the familiar is presented first. The results of the model's ability to predict the diffusion of prohibition using the adopter's physical location is rendered here. The analytical verification procedures indicate that the spatial dimensions are excellent predictors of the diffusion of prohibition, and that estimation error in the model is usually the result of specific identifiable historic incidents.

4.1 Analytical Prerequisites for Spatial Modelling

The model requires a pair of coordinates for each observation. In the spatial dimensions, these coordinates are county seat latitude and longitude. The location of the county seat was selected over the county centroid as the site of the adopter for several reasons. First, the county seat is usually near the centroid, and of the two, a named location is easier for subsequent researchers to identify with and relocate. Second, the county seat is usually at or near the population center of the county and, because the will of the populace is being measured, the county seat is a better choice than the physical center. Third, the county seat is the scene of the county's political power and activity, and as such, it was where the battle for adoption was most hardfought.
The longitude and latitude of each county seat, along with the county ICPSR code and county prohibition status in 1876, were passed to program1 which generated the connectivity network from the coordinate set. The output of program1 was read five times as input to program2. For each of the five runs, a different seed value was randomly selected for the random number generator routine in program2. The results of the program2 runs were submitted as input to program3. Program3 merged the results of the five independent runs to derive the model's omega pattern. With the derivation of the omega pattern, modeling ceased, and verification began.

4.2 Associating Predictions and Reality

The first phase of model verification involves identifying which of the 45 years in the study period had predictions made. The percentage of counties predicted dry in each column of the omega pattern is compared to the percentage observed dry in each year of the study period. Predictions are associated with reality, using the procedure outlined in the preceding chapter. As noted in the procedure's description not all years or predictions have an opposite to be matched with.

The omega pattern produced 13 generations that are assigned to observed years. The percentage of predicted dry counties in the omega pattern's first column corresponds best with the percentage observed dry in 1878. The percentage dry in 1881 was matched with the second column; 1885 with column three, 1887 with column four, 1891 with column five, and 1904 with column six. The columns for 1907, 1908, 1909, 1910, and 1916 are linked with columns seven through eleven,
respectively. Between 1916's prediction and 1917's prediction there are three unmatched columns in the omega pattern. 1917 is matched with column 15 and 1918 with column 18, leaving two more unmatched columns. The correlation between the 13 predicted-observed pairs is 0.99886 (Table 4.1).

Model accuracy varies, but the overall mean accuracy of the model is 71.6 percent. The accuracy for the first predicted year, 1878, is 96.2 percent. Accuracy for subsequent years declines at a steady rate until the predictions for 1904 when the model is 60.7 percent accurate. Between 1904 and 1907, accuracy drops another .6 percent. During the interval 1907 to 1916, accuracy begins to climb slowly but remains under 65 percent. For 1917, the predictive accuracy is 67.8 percent, and for 1918, the final year of simulation, it is 71.9 percent. The pattern of high initial accuracy followed by declining accuracy and then improvement during final predictions is common throughout all versions of the model and all scales of analysis. The ability of the model, using these two variables, to predict the growth of the movement in the range of the seventieth percentile speaks well for Hagerstrand's original choice of spatial dimensions for his model.

Just as the model's predictive accuracy is not constant over time it is not constant across space at any specific time. The overall accuracy of the model is the weighted mean accuracy of the model across time in each of the eight predefined regions. Local trends in accuracy, obliterated at the national level, are brought out when accuracy is examined regionally. The model is 65.94 percent accurate in New England, for example. In the remainder of the Northeast, the Mid-Atlantic region and the Northwest Territory states, the model
Table 4.1: Accuracy of the Model in the Spatial Dimensions

<table>
<thead>
<tr>
<th>Year</th>
<th>Observed Dry</th>
<th>Predicted Dry</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1876</td>
<td>12.0078</td>
<td>1</td>
<td>100.00</td>
</tr>
<tr>
<td>1877</td>
<td>13.3669</td>
<td>2</td>
<td>100.00</td>
</tr>
<tr>
<td>1878</td>
<td>13.1141</td>
<td>13.201</td>
<td>96.241</td>
</tr>
<tr>
<td>1879</td>
<td>13.3077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1880</td>
<td>13.5561</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1881</td>
<td>17.7032</td>
<td>16.502</td>
<td>88.810</td>
</tr>
<tr>
<td>1882</td>
<td>17.9429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1883</td>
<td>18.3934</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1884</td>
<td>19.3706</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1885</td>
<td>20.0669</td>
<td>19.934</td>
<td>84.299</td>
</tr>
<tr>
<td>1886</td>
<td>20.9354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1887</td>
<td>25.8608</td>
<td>24.158</td>
<td>78.484</td>
</tr>
<tr>
<td>1888</td>
<td>27.6704</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1889</td>
<td>28.8399</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1890</td>
<td>29.0509</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1891</td>
<td>29.3631</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1892</td>
<td>29.1007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1893</td>
<td>25.8238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1894</td>
<td>25.9729</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1895</td>
<td>23.9729</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1896</td>
<td>23.7143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1897</td>
<td>21.9843</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1898</td>
<td>21.9408</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1899</td>
<td>22.2341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1900</td>
<td>23.0879</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1901</td>
<td>23.0032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1902</td>
<td>24.5477</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1903</td>
<td>25.0885</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1904</td>
<td>37.5794</td>
<td>36.832</td>
<td>60.665</td>
</tr>
<tr>
<td>1905</td>
<td>37.6186</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>37.1347</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1907</td>
<td>41.2671</td>
<td>43.597</td>
<td>60.117</td>
</tr>
<tr>
<td>1908</td>
<td>47.7622</td>
<td>49.835</td>
<td>62.148</td>
</tr>
<tr>
<td>1909</td>
<td>56.7003</td>
<td>56.502</td>
<td>63.170</td>
</tr>
<tr>
<td>1910</td>
<td>61.2365</td>
<td>62.772</td>
<td>63.791</td>
</tr>
<tr>
<td>1911</td>
<td>64.8391</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1912</td>
<td>63.1489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1913</td>
<td>61.9357</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1914</td>
<td>59.4751</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>62.5892</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1916</td>
<td>69.4841</td>
<td>67.525</td>
<td>61.933</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71.056</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>73.564</td>
<td></td>
</tr>
</tbody>
</table>

1 No predictions are possible for 1876.
2 No predictions.
3 Multiple prediction for one year period.
### Table 4.1: Continued

<table>
<thead>
<tr>
<th>Year</th>
<th>Percent Observed Dry</th>
<th>Percent Predicted Dry</th>
<th>Percent Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1917</td>
<td>76.5541</td>
<td>77.294</td>
<td>67.850</td>
</tr>
<tr>
<td></td>
<td></td>
<td>78.878</td>
<td></td>
</tr>
<tr>
<td>1918</td>
<td>80.4590</td>
<td>80.627</td>
<td>71.983</td>
</tr>
<tr>
<td>1919</td>
<td>100.0000</td>
<td>100.000</td>
<td></td>
</tr>
</tbody>
</table>

sum 931.888  
mean 71.684

\[ r(\text{percent observed dry, percent predicted dry}) = 0.99886 \]

### Table 4.2: Accuracy of the Model in the Spatial Dimensions in Selected Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of Observations</th>
<th>Average Accuracy in Percent</th>
<th>Percent of Errors Type I</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td>67</td>
<td>65.940</td>
<td>94.93</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>153</td>
<td>58.770</td>
<td>99.46</td>
</tr>
<tr>
<td>Northwest Territory</td>
<td>436</td>
<td>66.530</td>
<td>93.23</td>
</tr>
<tr>
<td>Plains</td>
<td>618</td>
<td>69.267</td>
<td>44.42</td>
</tr>
<tr>
<td>South</td>
<td>988</td>
<td>75.507</td>
<td>36.68</td>
</tr>
<tr>
<td>N-S Border</td>
<td>369</td>
<td>75.503</td>
<td>38.71</td>
</tr>
<tr>
<td>Mountain West</td>
<td>266</td>
<td>84.995</td>
<td>0.0</td>
</tr>
<tr>
<td>Pacific West Coast</td>
<td>133</td>
<td>78.352</td>
<td>0.0</td>
</tr>
<tr>
<td>Total</td>
<td>3030</td>
<td>71.684</td>
<td>52.34</td>
</tr>
</tbody>
</table>
cent, respectively. The lowest regional accuracy of the model is for the Mid-Atlantic region. The predictive level for counties of the Great Plains region is 69.27 percent, almost the national average. Accuracy among Southern and Border counties is almost identical, 75.51 percent and 75.50 percent, respectively, and significantly above the average. The model is most accurate when predicting the status of counties in the Mountain West, 84.99 percent, and next most accurate, 78.35 percent, when making predictions for the Pacific West Coast region (Table 4.2). Most regional values are significantly different from the mean, provoking questions about the movement in each of these regions and about the performance of the model.

Just as the model varies in accuracy both over time and between regions, so does it also vary within regions over time. Regional variation in accuracy is a function of the actual distribution of dry counties in relation to the predicted distribution of dry counties over time. Spatial trends in accuracy are less well defined than temporal trends because they do not have an inherent pattern. The pattern of accuracy by regions can change with differing regional definitions. Further, the regional pattern is different depending on the variables used to construct the network. Regional variation cannot be construed, therefore, as a function of the extent of dry territory (number of counties dry) because the extent is matched in both the observed and predicted distributions during the initial phase of the analysis. The national, regional, and local differences in the observed and predicted distributions are true measures of the model's failings.

For analytical purposes, the predicted and observed distributions
are divided into eight regions. Model failings, and explanations for them, are presented in the next section using these regions as a frame of reference. Suffice it to say that the model's overall level of accuracy shows that the relationship between the observed and predicted growth patterns is a strong one. In some sections of the country at different times, however, the relationship is much stronger than at others.

4.3 Analysis of the Distribution of Predictive Errors

More important than knowing the number of errors is knowing the distribution of those errors in space and time. Analysis of error sites reveals trends in the diffusion of prohibition that are unexpected by the model, as well as trends that were previously unidentified in the literature. These error loci are identified in time and space, and spatial and historic explanations are offered for their presence.

4.31 Predictions for 1878

The distributions of dry counties in reality and in the invented world of the model for the year 1878, the first year of predictions, are very much like 1876, the base year for the model. Few changes have occurred in either the observed or the predicted distributions of dry territory since 1876. It is common for an innovation to spread slowly in the initial period of expansion or, in this case, re-expansion. The few newly dry counties are located in the Southern and Border regions (Fig. 4.1).

For 1878, model type I errors are located close to counties that
Figure 4.1

Results of Diffusion in the Spatial Dimensions, 1878

Figure 4.1

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
were dry in 1876 or counties that were dry in 1876 but that had reverted to the wet status by 1878. Type I errors appear as haloes around clusters of correctly predicted dry counties on the borders of New England, Iowa, Ohio, and the Indian Territory (Oklahoma). In these instances the model, expecting growth close to existing dry counties, has erroneously converted wet counties. On the Kansas-Indian Territory border an additional complication is present. Here, errors, rather than just erroneous prediction, reflect a problem of scale. The model predicts gradual county level adoption, but Plains states tended to adopt prohibition at the state level. A trend that causes a large proportion of overall model error. Predicted growth into the Mid-Atlantic region did not occur, in reality, because New York and Pennsylvania did not permit dry territory within their borders. This fact accounts for an additional substantial portion of model error during the simulation.

Many additional type I errors occur in the 1878 predictions adjacent to the scattered dry counties in the South and Border regions where local option was becoming the accepted method of enacting prohibition. In these type I error cases, the model selected an object county that did not actually adopt that year. Had another of the object counties of a dry subject county adopted, then, in that instance, the error would be random. If, on the other hand, no object county went dry, then the model overpredicts the local rate of growth (this may or may not mean that the regional rate of growth is being overpredicted). Given the scale of the model, it is difficult to determine which of these two cases is occurring in any given instance. An examination, however, of the state of Mississippi on the 1878 map
reveals at least two instances of probable random error (Fig. 4.1).

Type II errors occur wherever random error occurs, when counties go dry in isolation, and when local or regional growth of prohibition occurs at a rate greater than expected. No cases of isolated adoption exist at this early date. The rapid growth in Kentucky and East Tennessee, however, is under predicted. This rapid growth has been linked to a wave of fundamentalist Baptist fervor which passed through the region in the late 1870s and 1880s, and Methodist ministers who were permitted to speak out on civil matters after the Civil War (Whitner 1945, 19; Pearson and Hendricks 1967, 60; Farish 1938, 308-320). This instance of under prediction is brief; the actual growth rate slows during the next few years as the anti-drink, religious zeal wanes, but the predicted rate of increase remains stable and shortly catches up.

4.32 Predictions for 1881

In 1881, the second year of the study period that the model makes predictions for, another Plains state, Kansas, adopted prohibition. In the South there were 24 more dry counties than there had been in 1878. Four additional counties were dry in the border region. There was no growth in the number of dry counties in any of the other five regions.

Two basic problems that reduce the accuracy of this simulation throughout become apparent at this early stage. The first is a direct result of the model's primary assumption that wet counties adjacent to dry counties adopt in short order. The model's emulation of 1881, while partially reflecting the observed changes, makes many errors by
predicting growth of the movement in the Northeast where no growth occurred. In this region the movement was at a standstill. The model, following its rules of operation, however, predicts expansion in New York along the Vermont border, in Pennsylvania and Indiana along the Ohio border, and in several other locations (Fig. 4.2).

Why growth did not occur in the Northeast is a complex issue, but it can be related to several events. First, the Northeastern states had previous experience with prohibition during the 1850s. The dry territory in the Northeast in the last quarter of the 19th century was the vestige of that earlier period of expansion. The Northern states were loathe to support legislation that had failed them only thirty years previously. Second, recent European immigrants, of whom there was an ever increasing number, generally opposed prohibition. Finally, there were no great social movements in the North to link arms with the prohibition movement as there were in the South. These factors cannot explain why adoption did not occur at any one location. They were, nevertheless, influential throughout the region, to various degrees, during the course of the study.

The second problem is a result of the operational level of the model. In many instances, predicting adoption at the county level is appropriate, but in some cases, particularly in the Plains region, adoption occurs at the state level without a prerequisite adoption of a majority of the state's counties. Predictions for 1878 categorize the counties in the southern tier of Kansas as type I errors because the model incorrectly predicted them as adopters of prohibition. Adoption by the state of Kansas in 1881 reverses the relationship. Southern tier counties are correctly identified as dry in 1881 while
Results of Diffusion in the Spatial Dimensions, 1881

Figure 4.2

LEGEND
0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
the remainder of the state is incorrectly predicted wet. Over the next several predicted years, Kansas is progressively estimated drier.

In 1881 the model predicts that there are 86 dry southern counties. There were only 57, and the model only correctly identified 24 of these. The misprediction and overprediction creates substantial numbers of both type I and II errors. Many type II errors are clustered in southeastern Georgia far from any county that was dry five years previously. Adoption in this section of Georgia is assumed to be related to the activities of unknown fundamentalists who acted as change agents. Type I errors are scattered throughout the South and Border states; their only common feature is close proximity to dry counties.

In the remaining regions, the model fairly well reflects the patterns observed. Type II errors in these other regions are practically non-existent because the movement was at a standstill. There are no model errors in the Western regions. There are, however, a growing number of type I errors which appear throughout the remainder of the nation. These type I errors, which appear faintly as haloes around concentrations of dry territory in 1878, are now much more distinct (compare the patterns near Iowa, Ohio, and northern New England in Figs. 4.1 and 4.2). The model expects growth to have occurred in these places, but it did not.

4.33 Predictions for 1885, 1887, and 1891

The growth trends of the movement remained constant throughout the 1880s and into the early 1890s. Growth in the South remained rapid relative to the rest of the nation. Between 1881 and 1885, and
again between 1885 and 1887, the number of dry Southern counties
doubled. The number of dry counties in the border states climbed from
108 in 1881 to 129 in 1885, and from 201 in 1887 to 239 in 1891. West
Virginia, for example, was completely wet in 1885 and was over half
dry by 1887. In 1889, the states of North and South Dakota adopted
prohibition, reinforcing the tendency for Plains states to adopt at
the state level without any dry counties. Throughout the rest of the
nation, growth was negligible.

The model predicts prohibition status after 1881 and before the
turn of the century for the years 1885, 1887, and 1891 (Figs. 4.3-
4.5). Predictive errors by the model, while growing in number, are
extrapolations of previously defined failings; no new causes of system-
atic error are identified. Accuracy in the Northeast continues to
decline due to continued expansion of predicted dry territory in the
face of little change in the observed number dry counties. Western
accuracy remains at 100 percent until 1891 when the Pacific West Coast
region experiences its first adoption. Accuracy in the Pacific West
Coast region drops to 99.17 percent in 1891 because Modoc County,
California, over 1000 miles from any other dry county, adopted that
year. Accuracy in the Mountain West region remains at 100 percent
until after 1900.

Among the Plains states accuracy declines slightly between the
predictions for 1881 and 1887, but the number of correctly predicted
dry counties increased until 1889 because the model's emulation was
predicting the conversion of Kansas. The number of incorrectly pre-
dicted dry counties in Minnesota and Nebraska offset this trend, and
the result is an overall loss of accuracy in the region. Accuracy
Results of Diffusion in the Spatial Dimensions, 1885

Legend:
0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
Results of Diffusion in the Spatial Dimensions, 1887

Legend:
0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
Results of Diffusion in the Spatial Dimensions, 1891

LEGEND
0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
dropped sharply between 1887 and 1891, going from 77.13 percent to 57.93 percent due to the unexpected adoption in the Dakotas. Predictions for the South and Border regions drop in accuracy 21.1 percent and 28.5 percent, respectively, between 1881 and 1891. In these regions, the number of observed dry counties are very similar to the predicted number, but the model consistently mispredicts. The situation is worse in the South, where the total number of errors exceeds the number of correctly predicted dry counties by over three to one in 1887 and 1891. The errors encountered in the border region nearly equal the number of observed and predicted dry counties.

A fair proportion of the model errors in the South and Border regions can be classified as random, but systematic errors are present in much greater numbers. The random errors are predominantly located in states where the most rapid growth occurred: Mississippi, Alabama, Kentucky, and western Georgia. In north Texas, however, the model makes a serious systematic error by prematurely expanding dry territory there. Meanwhile, simulated adoption of coastal Georgia still has not begun even with the predictions for 1891, resulting in the continued presence of type II errors there. Finally, errors caused by statewide adoption in West Virginia are of short duration; predicted adoption was well under way there by 1891.

While predictions for the years 1885, 1887, and 1891 produce the same kinds of errors previously noted, the frequency at which they occur is much higher. The number of errors increased because the algorithm incompletely reflects reality, and its inability to replicate reality becomes more apparent with increasing lapse of time after 1876.
4.34 Predictions for 1904

1904 is the next year that the model predicts for in the spatial dimensions. The thirteen year hiatus in predictions is caused by an actual reduction in the number of dry counties during the late 1890s. The model could not predict this loss; it only predicts growth. 1904 was the first year after 1891 to show an increase in the percentage of dry counties over the percentage dry in 1891 (Table 4.1). Predictions, therefore, could not be associated with any of the intervening years.

Significant changes in the distribution of dry territory occurred in the years between 1891 and 1904 for four reasons. First, changes were due to an actual reduction in the number of dry counties during the mid 1890s. The vast majority of territorial losses occurred in the North where several states repealed their prohibitory laws; this substantially reduced the total number of dry counties. Second, changes in the distribution were caused by rapid growth of prohibition territory between 1900 and 1904 in the South and Border regions. Even during the 1890s, the movement was slowly expanding in the South. Third, the Supreme Court's decision that no state or local government could prevent the sale and distribution of liquor sold in its original package practically nullified dry status. In many places, prohibition laws were repealed because lawmakers felt that the laws had become superfluous. Fourth, distributional changes occurred because people's view of prohibition changed. Over a thirteen year period many people changed their minds about the effectiveness and utility of prohibition. There was also a considerable turnover in the voting population, which greatly affected decision-makers. Taken together, these
four factors represented a considerable influence on the distribution of observed dry territory. They had, unfortunately, no effect on the distribution of predicted dry territory, a fact that results in a much greater divergence between the two distributions than previously experienced.

Changes in how people felt about other, apparently unrelated, topics and changes in the demographic structure of a county directly and indirectly affected their view of prohibition. While many of these changes were highly localized, several were of national import. Continued immigration, the depression of 1896, the changed attitudes of young men returning from the Spanish-American War, and the activities of the Populist political movement in the South were all influential events. The destination of the vast majority of the immigrants was the Northeast, and their presence was influential in postponing adoption there and had some impact on accelerating repeal initiatives. The depression served to refocus activity and resources spent on prohibition in other directions, and redefined the moral stance of the nation (Clark 1933). The Populist movement in the South appears to have had a dual affect on prohibition. Populism promoted adoption in the South, yet, because of Populist stances on other issues, the alliance proved detrimental to prohibition in other regions.

Iowa, New Hampshire, Ohio, South Dakota, and Vermont repealed prohibition between 1891 and 1904. The status changes of these states from dry to wet resulted in massive concentrations of type I errors because the model ignores the possibility of recidivism. Ohio repealed prohibition in 1892, and only a few of its counties remained dry after 1900. New Hampshire and Vermont, both dry since before the
Civil War, repealed prohibition in 1903. The majority of Vermont's counties, however, remained dry, under local option until 1904 before rejecting prohibition also. The events in South Dakota seriously impaired the model's accuracy for 1891, but with the repeal of prohibition in 1897 accuracy in the Plains in 1904 improved by nearly one percent. This improvement occurred despite: the additional type I errors created on South Dakota's eastern border, continued predicted expansion in Minnesota and Nebraska, and the addition of wet counties in Iowa.

Iowa presents a unique case. The state remained dry throughout the entire study period, but starting in 1893, and continuing until 1915, the residents of a county could decide to permit liquor selling by means of local option elections (the reverse of local option in other states). If residents voted to allow saloons, a saloon keeper could set up shop, but he had to pay, in advance, the fine for running a saloon. After the advance payment, he had no further legal troubles and sold liquor as if he were in a wet county. The Iowa statute that established this system is called the Mulct Act.

Oklahoma, a border state, had wet counties for the first time in 1904. It was, however, omitted from the list of states that repealed prohibition between 1891 and 1904 for two reasons. Oklahoma was not yet a state in 1904, and more importantly, despite the presence of wet counties within its borders, the territory was dry. During the period from 1876 to 1889 Oklahoma was dry by federal mandate to protect the Indians from firewater. After Oklahoma was partially opened for white settlement in 1889, all reservation counties were dry, and all settlement counties were local option. In 1907 when Oklahoma was admitted
to the Union, the state's constitution contained a clause making the new state dry. Because of the status change from territory to state and the short duration of wet status in only some counties the unique events in Oklahoma are best presented separately.

Predictive errors for 1904 are much more numerous than in 1891 as evidenced by a drop of almost twelve percent in overall accuracy. The shifts in dry territory created a large number of additional errors throughout the North. Most of the errors throughout the remainder of the nation, however, result from continued misprediction, overprediction, and underprediction (Fig. 4.6).

In the Mid-Atlantic region there were no dry counties, but the model, since 1878, predicts an almost constant growth rate of fifty percent per prediction. By 1904, 46 of the region's 152 counties are incorrectly predicted dry. In the Mountain West region, where there were also no dry counties, the model remains completely accurate. Why are predictions for two like regions so different? The answer lies in the relative proximity of each region to dry territory. In 1876 the Mid-Atlantic bordered on two large concentrations of dry counties. The nearest dry counties to the West were located in the Indian Territory (Oklahoma) and Iowa. Simulated growth based on spatial proximity placed dry counties in Pennsylvania and New York after the first generation. If simulated growth had beelined for Colorado from Oklahoma, the minimal number of generations required for arrival would be four, but given the random selection of object counties for adoption the probable number of generations is many times that.

The growth of the movement in the South, given the retrenchment elsewhere, seems truly phenomenal. The reasons for the growth of the
Figure 4.6

Results of Diffusion in the Spatial Dimensions, 1904

Legend:
1. County Observed Wet and Predicted Wet
2. County Observed Dry and Predicted Wet (Type I Error)
3. County Observed Wet and Predicted Wet (Type II Error)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
movement in the South, and not elsewhere, lie in the Populist movement and the reactions to it by Southern conservatives. The Populists supported prohibition because it fit their overall views of how Americans should behave and be controlled. The opposition also favored prohibition, but as a mechanism to keep blacks in check and to disenfranchise them. Despite the statements of Southern conservatives, blacks also favored prohibition. The vast majority of counties that adopted prohibition in the between 1893 and 1897 were predominantly black. The mean percent black in 1890 for counties that adopted permanently in 1893 was 50.02 percent, in 1894 the percentage was 57.52, 69.98 percent in 1895, 54.75 percent in 1896, and 70.5 percent in 1897. When all parties favor an action, albeit for different reasons, it is generally taken.

Regarding predictions in the South for 1904, the number of type I errors decreased by 50 while the number of type II errors increased by 180. In 1904 there were 467 dry Southern counties, an increase over the number in 1891 (225). Many of the newly dry counties were located in western Florida, Virginia, and North Texas. Nearly all the Texan counties that were erroneously predicted dry in 1891, and many others, were really adopters by 1904. Mississippi, except for counties in the Delta region, was dry, as was much of Alabama, North Carolina, and Georgia.

Substantial concentrations of wet territory remained in the South, in southeast Texas, southern Louisiana, South Carolina, and along the Delta, for example. Southeastern Texas contained a large proportion of Catholic residents and was the site of most of the state's urban places at the turn of the century. South Louisiana
Residents were predominantly Catholics of French extraction, and southern Louisiana was the site of a major city, New Orleans. Both these factors, urbanization and Catholicism acted to slow the acceptance of the movement by the residents of these places. The Delta region, the band of counties paralleling the Mississippi River on either side from New Orleans to Cairo, Illinois, had to provide for travelers and rivermen. These two groups acted to make the Delta region different, more lawless if you will, than the remainder of the South, a difference that made Delta counties less likely to adopt.

The South Carolina government chose to establish a dispensary system rather than enact local option like its neighbors. In this way the people of South Carolina controlled the substance rather than banning it. Several counties in South Carolina did not have dispensaries and were therefore categorized as dry, but the vast majority did, and they were classed as wet. Rather than considering the lack of prohibition in South Carolina as resistance to the movement, it is better regarded as an alternative solution to the liquor problem whose presence the model ignores.

Regarding the Border states, growth occurred there slowly between 1891 and 1904. Most of Kentucky and West Virginia were dry by 1891. Missouri and Maryland were completely wet, in 1891 and 1904. In 1902 Tennessee passed legislation drying up all her rural counties. The model reflects these facts fairly well, but the reflection of 1904 in this region is the least accurate. The majority of the errors in the border region in 1904 are type II. They are concentrated in Tennessee, but they are also scattered throughout West Virginia and Kentucky. Type II errors in Tennessee are, once again, the result of the
model's failure to adapt to rapid change. Many type I errors are located in western Oklahoma territory and northern Missouri. The Oklahoma errors are transitory, and they disappear in 1907. The type I errors in Missouri are the result of the model expectation that counties in such close proximity to other dry counties (in Iowa) will adopt next.

Between 1893 and 1904 there was a radical change in the location of dry counties. The apparent relocation of dry territory into the South by means of expansion diffusion coupled with recidivism in the North brought about this shift. Both local and national trends shaped each individual decision to adopt. These trends cannot easily be measured, especially by a spatial diffusion model, even though they can be modeled with a good deal of accuracy by one. The major social trends and historical events obviously contributed significantly to many local decisions because of the spatial uniformity of the decisions within each region. On a larger scale, the repeal measures by state governments are also apparently correlated with the major social trends identified above.

4.36 Predictions for 1907, 1908, 1909 and 1910

The years 1907 through 1910 represent a continuation of the growth patterns identified in the analysis of 1904. Changes in the extent of dry territory during these four years demonstrate that the bandwagon was rolling as county after county jumped aboard. In 1907 41.27 percent of the counties were dry; by 1910 this figure had soared to 61.24 percent. The largest single year of increase during this four year period occurred between 1908, when 47.76 percent of all
counties were adopters, and 1909 when 56.7 percent were adopters.

Between 1904 and 1907 change in the overall pattern of dry territory was minimal (Fig. 4.7). Expansion continued in the South and Border regions as before. In the Northeast, only six additional counties adopted between 1904 and 1907, three of which were located in southern Illinois. Local option was enacted in Oregon in 1904, and three Oregon counties adopted in 1905. The West otherwise remained unchanged.

It was 1906 before local option elections were held in most Oregon counties, these elections determined status for 1907. After the 1906 elections, 29 of Oregon's 36 counties were dry. This was not a radical departure for Oregon, historically the state was the site of prohibition activity in the 1840s and was dry territory from 1844 to 1849. Oregon was not really part of the 'rough and tumble' West. Its earliest settlers founded missionary and agrarian colonies; their descendants were therefore somewhat more predisposed toward moralistic legislation than people in the remainder of the West. Model predictions for 1907 do not, of course, acknowledge the status change of Oregon, nor do the predictions for 1908, 1909, 1910, or even for 1918. According to the model, counties in Oregon are to remote for adoption to occur there.

In 1908 big changes begin to occur in the distribution of dry territory (Fig. 4.8). Colorado and New Mexico, enacted local option, and Georgia adopted statewide prohibition. Most of Georgia's counties were already dry, and it was the logical step to sanction the movement at the state level. This argument was frequently used by the Anti-Saloon League in their lobbying efforts for adoption at the next
Results of Diffusion in the Spatial Dimensions, 1908

Figure 4.8

Legend

0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
hierarchical level. Tennessee passed a law in 1908 prohibiting the
sale of alcohol in cities with populations under 50,000, thereby
limiting the number of wet counties in that state to four.

In 1909 the rate of change continues to accelerate, Alabama,
North Carolina, and Tennessee all adopted prohibition at the state
level. So many of their counties were already adopters, however, that
in the South there were only 56 additional dry counties in 1909 and 33
in the Border region. In each region about half of the newly dry
counties in 1909 were correctly predicted wet in 1908 and half were
represented as type II errors (Fig. 4.9).

All the states in the Northwest Territory region were much drier
in 1909 than they were in 1908. The number of dry counties in that
region climbed from 92 to 176. While most of these gains were made
along the region's southern border, growth in Michigan kept space by
adding ten new dry counties in 1908 and nineteen more in 1909. Fifty-
two type I errors in this region in 1908 were correctly predicted dry
in 1909.

By 1910 the movement had a foothold in every region. The Mid-
Atlantic, the final holdout, had dry counties in central Pennsylvania,
by that year, as a result of judicial decrees. Washington and Idaho
enacted local option in 1910. In South Carolina, only the coastal
counties still had dispensaries the remainder of her counties were dry
(Fig. 4.10).

What was the reason for the sudden rebirth of the movement out-
side the South after 1907? Several authors have related it to con-
stitutional adoption in Oklahoma in 1907 (Cherrington 1920b, 281-83;
Engleman 1975, 73). That event, it is believed, sparked a resur-
Results of Diffusion in the Spatial Dimensions, 1909

LEGEND

0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
Results of Diffusion in the Spatial Dimensions, 1910

LEGEND
0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
rection of hope in prohibitionists that inspired them to strive harder to complete their great task. The growing strength of the movement in the South also had much to do with the spread into other regions. Prohibition activity in the South was high pitched and some of the enthusiasm for the resurgence of high morals associated with it was felt in the adjoining regions. This is born out by the growth pattern which displays, to some degree, the characteristic wave form associated with spatial diffusion.

The development of an organization dedicated to expanding prohibition territory in a systematic manner also had much to do with the movement's resuscitation. Before this time there was no single organization devoted to organizing the movement politically at the national level without entering politics itself. The Prohibition party did not succeed because of the single issue campaigns of its candidates (they never succeeded in their fifty year history in electing a single major office holder). The WCTU was mostly a social club with franchises across the country, women had no part in the political arena. Together these forces shaped the growth of the movement before 1900. After 1900, however, the Anti-Saloon League took over as the driving force behind the movement. The ASL worked effectively as a change agent because it adhered to no lines, political or religious, but encouraged people to vote for prohibition either at apolitical referenda or to vote for candidates who had come out in favor of prohibition. This strategy enabled them to remain free of most political dealings and to sidestep the problems associated with being identified with a single candidate. They sought to make prohibition itself the candidate, and in carefully orchestrated campaigns they usually convinced most people
to vote for it.

On the whole, the simulated growth of the movement by the model is fairly accurate during this period even though the lowest accuracy encountered is recorded for 1907. The rapid rate of adoption between 1907 and 1910 counteracted much of the model's overpredictive propensity in earlier years in the Northeast. In New England, predictive accuracy is 41.79 percent in 1907; it continues to decline over the next four years as more dry counties are predicted in the region's wet southern portion. Accuracy also declines in the Mid-Atlantic states for the same reasons. In the Northwest Territory, the South, and the Border regions accuracy is on the upswing as these regions become saturated with both observed and predicted adopters. According to the model, adoption by a Western or northern Plains county is simply impossible. The isolated adoption in North Dakota and Oregon cannot be predicted until the model has erroneously (at this time) predicted all intervening counties dry. Isolated adoption, which was negligible in the early years, is proving itself the demise of the model during the second half of the study period.

4.37 Predictions for 1916, 1917, and 1918

The growth of the movement slowed after 1910, and it was not until 1916 that predictions were one again cast. In the South, only Louisiana, Arkansas, Texas, and Florida still contained wet counties in 1916, the other states had all adopted prohibition. Outside the South, Oregon, Idaho, Iowa (again), Colorado, Washington, and Arizona all adopted between 1910 and 1916. Nebraska, Arkansas, South Dakota, and Utah adopted in 1917. In the Northeast, Michigan (1917) was the
sole adopter prior to 1918.

Model accuracy in 1916 is down slightly from the 1910 level, but it is at 67.85 percent in 1917; the highest it has been since 1891. The major reason for the decline in 1916 is the model's failure to identify adoption by Western counties. In the Mountain West, regional accuracy drops from 85.37% in 1910 to 43.5% in 1916, and in the Pacific West from 65.42% to 42.11%. Other problems in prediction result from the continued overprediction of dry counties in the Mid-Atlantic region (Fig. 4.11).

Predictive accuracy for 1917 is little different from accuracy levels noted in 1916 except that, on the whole there are fewer errors. Errors in the South, Plains, and Border States decrease as the model finishes converting counties in the dry states there. Accuracy in the South climbs from 77.84 percent in 1916 to 85.38 percent in 1917. In the Plains, accuracy increased 7.1 percent from 63.2 percent in 1916. Among the Border states accuracy reaches 91.87 percent in 1917 up almost six percent from 1916 (Fig 4.12).

Throughout the remainder of the nation accuracy is low in 1917. In New England the model is 43.28 percent accurate in 1917 down 1.5 percent from 1916. The Mid-Atlantic region is only 13.725 percent accurate, the lowest accuracy provided by the model for any year or region. The Northwest Territory region's accuracy is up 11.9 percent from 1916 and stands at 50.45 percent. In both Western regions accuracy is between 41 and 42 percent in 1917.

The adoption of prohibition by 28 states by 1918 and by many of the counties outside these states rendered 80.46 percent of all counties dry. The model predicts that 80.63 percent of all counties are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Results of Diffusion in the Spatial Dimensions, 1916

LEGEND
0 County Observed Wet and Predicted Wet
1 County Observed Dry and Predicted Dry
2 County Observed Wet and Predicted Dry (Type I Error)
3 County Observed Dry and Predicted Wet (Type II Error)
Figure 4.12

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
dry in 1918, even though only four Western counties were predicted
dry. The minimum percent that could be predicted correctly for 1918
is 61.09, but the actual percentage correctly predicted is 71.98
percent. The largest regional errors occur in the Mid-Atlantic, the
Northwest Territory, and the two Western regions. Mid-Atlantic accu-
ricy is the lowest observed for 1918, 18.95 percent. In the Western
regions, the model is 42.52 percent accurate in the Mountain West and
39.1 percent accurate in the Pacific West. In the Northwest region it
is 55.73 percent correct. In the two eastern regions errors are again
due to overprediction, and in the Western regions errors are due to
underprediction (Fig 4.13).

Accuracy in the remaining regions is higher than the mean in
1918. In New England accuracy is at 74.63 percent due to the re-
adoption by Vermont, New Hampshire, and Massachusetts. The failure of
the model to predict adoption in the Dakotas, and its prediction of
Minnesota as dry, hinder accuracy in the Plains, but the percentage of
correct predictions remains above the mean at 76.66 percent. In the
South and Border regions where observed adoption was practically
universal by 1918, the model is 88.75 percent and 93.5 percent accu-
rate, respectively.

The inability of the model to predict status change in the United
States west of the Rockies, and in the northern Plains, is primarily a
function of distance and the number of intervening opportunities for
contact and adoption. The Western counties are far from the original,
1876, sites of dry counties. It would take, for example, seven pre-
dictive generations, minimally, for a conversion signal from northwest
Iowa to reach, via intermediary counties in South Dakota and Minne-
sota, southeast North Dakota. In actuality it took eleven. If it is assumed that the probability of contact between each subject county and one of its object counties is 0.083 (1/12), which it is not because of weighting for distance, then the probability of contact occurring at a specific county seven counties away from a dry subject county in seven generations is 0.0000000279. Contact in eleven generations, however, is much more probable. Because there are several adjacent counties in northwest Iowa that can transmit the innovation initially, and each new adopter can transmit after contact. With the number of dry counties, and the potential number of choices by each, contacts are made in every direction in an almost uniform pattern (Yuill 1964).

A second potential problem crops up when the predictive diffusion wave front generated from the Indian Territory approaches the Mountain states of the West. Examination reveals a sudden and drastic decrease in county density in the Western states when compared to the Plains, Southern, and Border states. Any random number drawn for a predicted dry county on the western border of a Plains state is much more likely to fall within the probability range of a sector pointing in the eastern direction than in the western direction because of how probability ranges are assigned. No examples of this phenomenon can be found because the model simply does not have ample opportunity to expand into the West before the criteria for the onset of national prohibition are achieved.

One could easily provide a scenario where the events described above could occur. Suppose, for example, that there is a predicted dry county on the Texas side of the Texas-New Mexico border. The scanner identifies it and draws a random number for it. The close
evenly spaced Texas counties that surround this county on six sectors have high, relatively even, probabilities of contact. The nearest New Mexican county has a much lower probability owing to its greater distance. Further, because of the irregular pattern of counties in New Mexico, one or more of the sectors that face into New Mexico are empty. As a result, the probability of contacting a county in New Mexico is greatly reduced, the overall effect of the density shift is that of a permeable barrier.

By 1918, the counties of the states east of the Rockies have been converted, with few exceptions, to the dry status (Fig. 4.13). Only extremely isolated counties have not been predicted dry, as in peninsular Florida, south Texas, the western portions of the Dakotas, and northern Minnesota. There still are a large number of type I errors in the Northeast because of the observed slowness of adoption in the region. New Jersey is without a dry county. Wisconsin and much of Illinois, for example, although predicted dry, were wet in 1918. In the West, the same pattern of type II errors and correct predictions of wet counties, first encountered in 1909, continues.

The errors in the last three years of prediction decline primarily because of saturation effects, but these effects are not to be dismissed lightly. Saturation of the system is an important and expected event in diffusion modelling and research. As was mentioned, the model is substantially more accurate than it minimally could be in 1918. This is despite two glaring, extensive cases and several minor cases of systematic error. Whatever else the model shows, it has demonstrated that the prohibition movement diffused, to a large degree, spatially.
4.4 Summary

The verification procedures used in the analysis demonstrate that the spatial diffusion model predicts the diffusion of prohibition very well. Model errors, for the most part, represent instances where the innovation did not spread in a purely contagious fashion. The model is incorrect for only 28 percent of all predictions. These errors are not randomly distributed in space and time, instead they form provocative patterns. Detailed examination of these spatial and temporal patterns provide a great deal of information regarding both the model, its failings, and the prohibition movement.

In 1876 most of the dry counties were located in the North, particularly in the Mid-West and New England. By the middle years of the study period, however, many of these counties had abandoned prohibition, resulting in a large number of type I errors in the region. The remainder of the North was the site of many other type I errors because counties there did not adopt until much later than they were predicted to have. It is surmised that the lateness of adoption is related to previous experience with prohibition in the region in the 1850s. The observed recidivism is the result of a lag in regional disgruntlement with the innovation. Factors such as the increasing proportion of immigrants in the region, the depression of 1896, the original package law, and attributes of Southern adoption also aided the evacuation of the movement from this region.

In the South the reverse is true. From 1876 onward the movement, which had never been strong there before, grew in the South. The growth observed in the South is well predicted relative to the predictive ability of the model in other regions. At its worst the model
is 54.92 percent accurate in the South. Much of the error in the South is random or the result of short term over or underprediction. Historically the unprecedented growth in the South is related to a wave of fundamentalist religious fervor that swept the region in the late 1870s and 1880s, the new ability of Southern churchmen to intervene in civil affairs, the Populist political movement in the 1890s, and to the activities of the Anti-Saloon League thereafter.

Most of the model errors in the Plains region resulted from the adoption by states instead of counties in this region. While the model plodded along converting a tier of counties per year entire blocks went dry. Kansas went dry in 1881, North and South Dakota in 1889 (South Dakota repealed in 1896), Iowa and Oklahoma (then Indian Territory) were dry from the beginning, although each had brief periods of dampness resulting in recidivous errors. On the other side, much of Minnesota, Nebraska, and South Dakota remained wet throughout the study period, but each cycle of predictions represented them as progressively drier. Despite these errors caused, in part, by scale differentials and the errors caused by the isolated adoption of North Dakota, regional model accuracy remained above 51 percent throughout the study period.

Model errors in the Western regions do not occur before 1891 and are of little consequence before 1907. In the Pacific West, accuracy drops to 76.34 percent from 98.4 percent in 1904. In the Mountain West accuracy is 43.5 percent in 1916 down from 94.66 percent in 1910. Western errors are always type II revealing the extent of underprediction in this region. The model does not make predictions in the West before 1917, and it only predicts four dry counties in 1918 in the two
regions. The reason for this failure is the great distance between any dry county in 1876 and the Western counties.

Throughout this chapter explanations were offered for the failure of the model to predict accurately. These explanations fell into two major categories, random error and systematic error. Random error is expected to occur when using any simulation routine, but systematic error, when identified, is both a source of information and a reason for concern. As an information source, systematic error says that some events are outside the explanatory capacity of the model. As a source of concern it says that the model does not consider certain, obviously important, pieces of data.

Of the many sources of systematic error unearthed in this analysis they may all be classified into a very few categories. Error resulting from recidivism and error resulting from adoption by state units cannot be eradicated because of the definition of adoptive units and what constitutes adoption. Error resulting from overprediction is encountered when the model predicts adoption where it has not yet occurred. Underprediction is just the opposite. In each case these errors are the result of movement growth modified by aspatial variables. The final type of systematic error, isolated adoption, is a special case of underprediction associated with relocation diffusion. Each occurrence of systematic error is unique because the people and events involved are unique, but their commonalities bind them into one of these types.

The expansion of the model's capacity to explain events by requiring it to consider other pieces of information is necessarily the next step in the simulation modeling of the diffusion of prohi-
bition. Towards this goal a new version of the model, using an alternate coordinate system, is discussed in the next chapter.
CHAPTER FIVE

ANALYSIS OF SIMULATED DIFFUSION IN THE URBAN-ETHNIC DIMENSIONS

As in chapter four, the ability of the diffusion model to predict the paths of the adoption of prohibition is tested. Here, however, the coordinates used to generate the distance and probability networks are socio-economic characteristics of the county. The two variables required as input dimensions to the model are distilled from 17 variables by factor analysis, selected from the censuses of 1900 through 1920. The basic assumption of this model variant is that prohibition diffused in the United States from a group of similar counties to counties that were increasingly less like the original adopters. The model assumes that counties with similar socio-economic characteristics have similar goals, needs, and outlooks, and that they are prone to act and react to the same stimuli in a similar fashion. The operational explanation for adoption is, therefore, that an object county adopted because it is like a subject county.

This variant two-dimensional model is not as accurate as the spatial variant, but it provides direct information regarding the nature of adopters not derived from the spatial analysis making it also valuable.

5.1 Building the Coordinate System

While the spatial model uses a made-to-order coordinate system for its dimensions, no such convenient coordinate system exists for the dimensions of the social model. Rather, a new, synthetic set of dimensions was constructed from 17 socio-economic indicators. Factor
analysis reduced the number of variables from 17 to two, the number of data dimensions required by the simulation model.

Factor analysis reduces the number of variables, by identifying on the basis of correlation between variables any underlying regularity in the data. The first step in factor analysis is the calculation of the correlation matrix for the variables in question. From the correlations, a factor pattern is constructed to represent the degree of similarity observed across correlations. The degree of similarity is a mathematical transformation (represented by a set of linear combinations) of the original data. The first linear combination in the factor pattern accounts for the largest proportion of the variance explainable by factor analysis techniques. The second linear combination accounts for the next largest amount of the residual variance. Ideally, a linear combination is calculated for each variable included in the analysis, but the latter combinations frequently do not have strong contributions from any variables, leaving them to explain miniscule proportions of the variance observed.

The values of the linear combinations are rotated (mathematically adjusted to simplify the factor pattern) to make the variance accounted for by each orthogonal. Orthogonality adjusts the values of linear combinations making each independent. In the process, adjustments are made to the amount of variance explained by each linear combination. These adjustments result in a more equal splitting of the explained variance while maintaining the same overall amount of variance explained.

The factor analysis routine used is the maximum likelihood option coupled with a varimax rotation from the Statistical Analysis System's
Proc Factor program (SAS 1982, 309-45). Scoring coefficients were also generated by proc factor. The scoring coefficients and the raw data were then subjected to SAS's Score procedure, which generated the two new variables FACTOR1 and FACTOR2 (SAS 1982, 485-92).

The 17 social, economic, and demographic variables included in the factor analysis were drawn from the censuses of 1900, 1910, and 1920 and the Census of Religious Bodies for 1896, 1906, 1916, and 1926. Despite intentions to include data from earlier censuses, the number of observations had already been reduced from 3030 to 2499 by the inclusion of the 1896 Census of Religious Bodies. These 531 observations were lost because these counties came into existence after 1896. The states of Oklahoma, Arizona, and New Mexico are lost to this analysis for this reason. Inclusion of data from earlier censuses would have removed too many more observations from the analysis.

The variables and variable names used for each county are: total population 1920, TP0P920; total population 1910, TP0P910; total population 1900, TP0P900; percent foreign born 1920, FBRN920, percent foreign born 1910, FBRN910; percent foreign born 1900, FBRN900; percent Negro 1920, NEGR920; percent Negro 1910, NEGR910; percent Negro 1900, NEGR900; percent employed in manufacturing 1920, WORK920; percent employed in manufacturing 1900, WORK900; percent urban 1910, URB910; percent urban 1900, URB900; population density 1910, DENS1910; percent Roman Catholic 1920, PCATH920; percent Roman Catholic 1910, PCATH910; and percent Roman Catholic 1900, PCATH900. Percentages and population densities were calculated from the raw census figures. The percent Roman Catholic was interpolated from the Censuses of Religious
Bodies for 1896, 1906, 1916, and 1926 and the appropriate Decennial Censuses.

The way these variables loaded on each of the factors indicates the presence of an underlying structure amongst the variables used in the analysis (Table 5.1). A factor loading is a coefficient for a variable identifying its relation to the factor. When an orthogonal rotation has been applied, the loading coefficient of a variable is equivalent the correlation coefficient between the factor and the variable. A variable is considered an important contributor to a factor, if its factor-pattern score (loading coefficient) is greater than .5. FACTOR1 represents the level of urbanization (urbanness) of each county and accounts for 30.101 percent of the variance identified within the 17 variables. The variables that loaded strongly on FACTOR1 are: TP0P920, TP0P910, TP0P900, WORK920, WORK900, URB910, URB900, and DENS910.

FACTOR2, accounting for 26.002 percent of the variance, represents the ethnic, racial, and religious composition (ethnicity) of the county. The percentages of foreign born persons, Negroes, and Roman Catholics loaded on FACTOR2, but the foreign born and Roman Catholic variables loaded negatively, while the Negro percentage loaded positively. This arrangement construes ethnicity as a continuum from counties with Negroes, Protestants, and native born persons on the positive side of FACTOR2 scores, and foreign born persons, Roman Catholics, and other races concentrated on negative side of FACTOR2 scores. The variable FBRN900 did not load into either factor at an acceptable level as its loading on each factor is less than .5. Both factors combined account for 56.106 percent of the variance observed.
Table 5.1: Orthogonal Rotation of Factor Pattern and Standardized Scoring Coefficients (n=2499)

<table>
<thead>
<tr>
<th>Variable</th>
<th>FACTOR1</th>
<th>FACTOR2</th>
<th>FACTOR1</th>
<th>FACTOR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPOP920</td>
<td>.86013</td>
<td>.06958</td>
<td>.18595</td>
<td>.07042</td>
</tr>
<tr>
<td>FBRN920</td>
<td>.35241</td>
<td>-.73004</td>
<td>.02909</td>
<td>-.15660</td>
</tr>
<tr>
<td>NEGR920</td>
<td>.08099</td>
<td>.80576</td>
<td>.06714</td>
<td>.20203</td>
</tr>
<tr>
<td>WORK920</td>
<td>.59460</td>
<td>-.20418</td>
<td>.11288</td>
<td>-.01300</td>
</tr>
<tr>
<td>PCATH920</td>
<td>.36674</td>
<td>-.61588</td>
<td>.03800</td>
<td>-.12780</td>
</tr>
<tr>
<td>TPOP910</td>
<td>.87192</td>
<td>.09511</td>
<td>.19033</td>
<td>.07739</td>
</tr>
<tr>
<td>FBRN910</td>
<td>.32081</td>
<td>-.74057</td>
<td>.02176</td>
<td>-.16114</td>
</tr>
<tr>
<td>NEGR910</td>
<td>.15151</td>
<td>.78092</td>
<td>.08048</td>
<td>.20033</td>
</tr>
<tr>
<td>URB910</td>
<td>.57670</td>
<td>-.33463</td>
<td>.10100</td>
<td>-.04601</td>
</tr>
<tr>
<td>DENS910</td>
<td>.65598</td>
<td>.16377</td>
<td>.14869</td>
<td>.08077</td>
</tr>
<tr>
<td>PCATH910</td>
<td>.39552</td>
<td>-.62740</td>
<td>.04456</td>
<td>-.12883</td>
</tr>
<tr>
<td>TPOP900</td>
<td>.87189</td>
<td>.10077</td>
<td>.19037</td>
<td>.07878</td>
</tr>
<tr>
<td>FBRN900</td>
<td>.33955</td>
<td>-.14175</td>
<td>.06257</td>
<td>-.01489</td>
</tr>
<tr>
<td>NEGR900</td>
<td>.07719</td>
<td>.80688</td>
<td>.06640</td>
<td>.20207</td>
</tr>
<tr>
<td>WORK900</td>
<td>.67247</td>
<td>-.21581</td>
<td>.12860</td>
<td>-.01101</td>
</tr>
<tr>
<td>URB900</td>
<td>.68345</td>
<td>-.27614</td>
<td>.12718</td>
<td>-.02507</td>
</tr>
<tr>
<td>PCATH900</td>
<td>.39109</td>
<td>-.55391</td>
<td>.04819</td>
<td>-.11114</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvalue FACTOR1</th>
<th>6.114786</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalue FACTOR2</td>
<td>3.423275</td>
</tr>
<tr>
<td>Post Rotation Eigenvalue FACTOR1</td>
<td>5.117782</td>
</tr>
<tr>
<td>Post Rotation Eigenvalue FACTOR2</td>
<td>4.420280</td>
</tr>
<tr>
<td>Amount of variance explained</td>
<td>9.538062</td>
</tr>
</tbody>
</table>

| Percent of Variance Explained | 56.1063 |
| Percent Explained by FACTOR1 | 30.1046 |
| Percent Explained by FACTOR2 | 26.0017 |
within the 17 variables.

The actual creation of the coordinates of each observation for the diffusion model is performed by the scoring procedure. The standardized scoring coefficients for each variable for the two factors are calculated from the loadings based on regression techniques in that the scores are beta coefficients (Table 5.1). The z-score of each observation for each variable is multiplied by its respective scoring coefficient for each of the two kept linear combinations. The products for all multiplications for each observation are then summed. The results for each observation are interpreted in three ways. First, the two values are that observation's FACTOR1 and FACTOR2 scores. Second, they are that observation's position relative to the mean score of all observations for each of the two linear combinations, and third, they are that observation's coordinates in the two dimensions (variable names FACTOR1 and FACTOR2) used in the current version of the diffusion model.

The expected distribution of data points in the two factor system is bivariate normal. The mean of the factor scores for a factor is zero and the variance is, ideally, one, but the computed scores only estimate the factor. A county whose score for FACTOR1 and for FACTOR2 equals (1,1) is one positive standard deviation from the mean of each factor. Of all counties 68 percent are expected to be located within a circular area centered on (0,0) and one standard deviation in radius. A county's score on the two factors is a reflection of what that county was like relative to all other counties. The higher the ethnicity score the higher percentage of a county's residents were Negro, and the lower the urbanness score the more rural a county was.
Unlike the spatial dimensions some orientation in the distribution of counties in this artificial coordinate system is required. It is generally known how counties in Kansas differ spatially from counties in California or Maine, but how much does a county at (0,0) differ from a county at (1,1) or (-1,-1) in this system? What are each of these counties like in terms of the variables used to create the coordinates? What are the ranges of values in each dimension? These questions are addressed in the following section.

5.2 The Distribution of Counties in the Social Dimension

FACTOR1 scores, which define urbanness, range from -1.01, Rio Blanco county, Colorado, to 21.21, New York County, New York. FACTOR2 scores range from -4.41, Presidio County, Texas, to 6.54, New York County. Dividing the coordinate system into four quadrants permits easy identification of score patterns and facilitates the pending analysis. Quadrant one (QUAD1), where county scores in both dimension are positive, holds 13.6 percent of the 2499 counties included in this aspect of the analysis. When both factor scores are positive the county is more urban and has a higher percentage of Negroes and Protestants than the average county. Quadrant two (QUAD2), where FACTOR1 is positive and FACTOR2 is negative, houses counties that are again more urban than the mean county, but with a higher percentage of foreign born and Roman Catholic residents, contains 21.6 percent of the 2499 counties. Quadrant three (QUAD3), where both factors are negative, indicating counties that are rural and have higher than average concentrations of foreign born and Roman Catholic residents, holds 31.1 percent of the counties. Within quadrant four (QUAD4), the
opposite of QUAD2, reside the remaining 33.7 percent of the counties.

What are the socio-economic variable values of counties that score at various points in the distribution? To answer this question counties scoring at three points in the distribution have been selected for discussion (Table 5.2). The county that scored closest to the origin (FACT0R1= -.0015, FACT0R2= .0689) is Lawrence County, Indiana. Its variable values for 1910 are: TPOP910= 30,625, FBRN910= 2.7 percent, NEGR910= 1.1 percent, DENS910= 67.16 people per square mile, URB910= 39.7 percent and CATH910= 2.63 percent. Glynn County, Georgia, scored closest to (1,1) with the scores (FACT0R1= 1.0128, FACT0R2= 1.4196). Glynn County has the following values for its 1910 variables: TPOP910= 15,720, FBRN910= 2.4 percent, NEGR910= 62.2 percent, DENS910= 13.02 persons per square mile, URB910= 64.8 percent, and CATH910= 1.48 percent. Ferry County, Washington, scored -.8062 on FACT0R1 and -1.0781 on FACT0R2, making it the county scoring closest to (-1,-1). Ferry County's variable values for 1910 are: TPOP910= 4,800, FBRN910= 33.8 percent, NEGR910= .2 percent, URB910= 0 percent DENS910= 2.16 persons per square mile, and CATH910= 9.21 percent.

From these variable values for these three counties, the pattern of scores on the distribution is readily comprehended, especially along FACT0R2, from the values observed for the variables that contribute most heavily to each factor. Lawrence was composed of native born, white non-catholics, it therefore scored close to zero. Glynn was like Lawrence except for its large black population, it therefore scored positively. One of three Ferry County residents were foreign born and one in ten was a Catholic, it therefore scored negatively. Care must be taken, however, lest too much be assumed. If, for exam-
Table 5.2: Selected Variables and Factor Scores for Selected Counties

<table>
<thead>
<tr>
<th>Variable</th>
<th>Glynn Co.</th>
<th>Lawrence Co.</th>
<th>Ferry Co.</th>
<th>Chester Co.</th>
<th>Webb Co.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Georgia</td>
<td>Indiana</td>
<td>Washington</td>
<td>Pennsylvania</td>
<td>Texas</td>
</tr>
<tr>
<td>FACTOR1</td>
<td>1.0128</td>
<td>0.0015</td>
<td>-0.8062</td>
<td>1.0257</td>
<td>1.3497</td>
</tr>
<tr>
<td>FACTOR2</td>
<td>1.4196</td>
<td>-0.0689</td>
<td>-1.0781</td>
<td>-0.1659</td>
<td>-4.2799</td>
</tr>
<tr>
<td>TPOP920</td>
<td>19,378</td>
<td>32,822</td>
<td>5,143</td>
<td>115,120</td>
<td>29,152</td>
</tr>
<tr>
<td>FBRN920</td>
<td>2.0%</td>
<td>1.4%</td>
<td>28.1%</td>
<td>8.4%</td>
<td>53.5%</td>
</tr>
<tr>
<td>NEGR920</td>
<td>50.9%</td>
<td>0.7%</td>
<td>0.0%</td>
<td>9.9%</td>
<td>0.1%</td>
</tr>
<tr>
<td>WORK920</td>
<td>9.0%</td>
<td>9.2%</td>
<td>1.3%</td>
<td>10.9%</td>
<td>1.0%</td>
</tr>
<tr>
<td>PCATH920</td>
<td>1.4%</td>
<td>2.7%</td>
<td>7.5%</td>
<td>10.8%</td>
<td>64.8%</td>
</tr>
<tr>
<td>TPOP910</td>
<td>15,720</td>
<td>30,625</td>
<td>4,800</td>
<td>109,213</td>
<td>22,503</td>
</tr>
<tr>
<td>DENS910*</td>
<td>35.8</td>
<td>67.16</td>
<td>2.2</td>
<td>140.6</td>
<td>7.0</td>
</tr>
<tr>
<td>FBRN910</td>
<td>2.4%</td>
<td>2.7%</td>
<td>33.8%</td>
<td>9.8%</td>
<td>48.7%</td>
</tr>
<tr>
<td>NEGR910</td>
<td>62.2%</td>
<td>1.1%</td>
<td>0.2%</td>
<td>9.7%</td>
<td>0.2%</td>
</tr>
<tr>
<td>URB910</td>
<td>64.8%</td>
<td>39.7%</td>
<td>0.0%</td>
<td>38.8%</td>
<td>66.0%</td>
</tr>
<tr>
<td>PCATH910</td>
<td>1.8%</td>
<td>4.5%</td>
<td>18.6%</td>
<td>13.0%</td>
<td>77.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Colorado</td>
<td>New York</td>
<td>Tennessee</td>
<td>Washington</td>
<td>Georgia</td>
</tr>
<tr>
<td>FACTOR1</td>
<td>-1.0186</td>
<td>0.0048</td>
<td>-0.4229</td>
<td>0.9947</td>
<td>0.0072</td>
</tr>
<tr>
<td>FACTOR2</td>
<td>-0.0490</td>
<td>-0.8592</td>
<td>0.9915</td>
<td>-1.0645</td>
<td>1.0307</td>
</tr>
<tr>
<td>TPOP920</td>
<td>3,135</td>
<td>31,871</td>
<td>5,996</td>
<td>44,745</td>
<td>16,362</td>
</tr>
<tr>
<td>FBRN920</td>
<td>4.1%</td>
<td>8.1%</td>
<td>0.0%</td>
<td>25.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td>NEGR920</td>
<td>0.2%</td>
<td>0.1%</td>
<td>28.6%</td>
<td>0.2%</td>
<td>71.0%</td>
</tr>
<tr>
<td>WORK920</td>
<td>0.6%</td>
<td>4.8%</td>
<td>1.4%</td>
<td>23.3%</td>
<td>0.7%</td>
</tr>
<tr>
<td>PCATH920</td>
<td>1.8%</td>
<td>23.4%</td>
<td>0.0%</td>
<td>11.7%</td>
<td>0.0%</td>
</tr>
<tr>
<td>TPOP910</td>
<td>2,332</td>
<td>33,458</td>
<td>5,874</td>
<td>35,590</td>
<td>16,552</td>
</tr>
<tr>
<td>DENS910*</td>
<td>0.7</td>
<td>18.2</td>
<td>55.4</td>
<td>18.5</td>
<td>51.6</td>
</tr>
<tr>
<td>FBRN910</td>
<td>8.2%</td>
<td>11.5%</td>
<td>0.1%</td>
<td>30.9%</td>
<td>0.1%</td>
</tr>
<tr>
<td>NEGR910</td>
<td>0.3%</td>
<td>0.2%</td>
<td>30.3%</td>
<td>0.2%</td>
<td>69.4%</td>
</tr>
<tr>
<td>URB910</td>
<td>0.0%</td>
<td>3.2%</td>
<td>0.0%</td>
<td>61.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>PCATH910</td>
<td>1.3%</td>
<td>10.8%</td>
<td>0.0%</td>
<td>9.3%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

---

*Density is in persons per square mile.

FACTOR1

| 5 | 9 | 4-A | 1 |

FACTOR2

| 8 |

| 3 | 6 |
ple, a county has both high proportions of black and foreign born residents, then its ethnicity score may be near zero because the two will counteract each other.

The pattern of values of the variables contributing to FACTOR1, however, leaves a hazy notion of the determining criteria for FACTOR1 scores. Glynn County scored approximately one standard deviation higher on FACTOR1 than did Lawrence County with a smaller total population and a lower population density, but compared to Lawrence County Glynn County had nearly twice the percentage of its population located in an urban place. Given the weighting scores, .19 for TPOP910 and 0.149 for DENS910, the explanation for the higher Glynn County urbanness value is that its density values are so far above the mean that they push the counties value well beyond that of a more highly populated, although less dense Lawrence County. Even though URB910 loaded less strongly than did the other two variables, it is in this example the best indicator of direction along FACTOR1. If two counties have identical values for the variables that load strongly on FACTOR1 the county that scores absolutely higher on FACTOR2 has the higher FACTOR1 value because all 17 variables contribute to the factor not just those that load intently.

Within the socio-economic dimensions, an underlying spatial pattern of counties is readily discernable. Regions, particularly those that are compact longitudinally, cluster well. The tendency of each region to cluster provides a key to orienting oneself to the distribution. More importantly, it provides evidence of the linkages between county similarity measured by these two variables and county propinquity (Table 5.3).
Table 5.3: Categorized County Factor Scores by Selected Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>QUAD1</th>
<th>QUAD2</th>
<th>QUAD3</th>
<th>QUAD4</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td>1</td>
<td>51***</td>
<td>11</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>8</td>
<td>115***</td>
<td>17</td>
<td>8</td>
<td>148</td>
</tr>
<tr>
<td>NW Territory</td>
<td>8</td>
<td>176*</td>
<td>166*</td>
<td>65</td>
<td>415</td>
</tr>
<tr>
<td>Plains</td>
<td>66</td>
<td>345***</td>
<td>15</td>
<td></td>
<td>426</td>
</tr>
<tr>
<td>South</td>
<td>289*</td>
<td>25</td>
<td>44</td>
<td>465**</td>
<td>823</td>
</tr>
<tr>
<td>N-S Border</td>
<td>35</td>
<td>22</td>
<td>36</td>
<td>301***</td>
<td>394</td>
</tr>
<tr>
<td>Mountain West</td>
<td>26</td>
<td>95***</td>
<td>1</td>
<td></td>
<td>122</td>
</tr>
<tr>
<td>Pacific Coast</td>
<td>42*</td>
<td>66**</td>
<td></td>
<td></td>
<td>108</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>341</td>
<td>525</td>
<td>780</td>
<td>855</td>
<td>2499</td>
</tr>
</tbody>
</table>

*** Cell represents over 75 percent of regional observations.
** Cell represents over 50 percent of regional observations.
* Cell represents over 25 percent of regional observations.

URBAN

QUAD2

(FACTOR1 +, FACTOR2 -)

FOREIGN BORN & CATHOLIC

QUAD3

(FACTOR1 -, FACTOR2 -)

NEGRO

QUAD1

(FACTOR1 +, FACTOR2 +)

QUAD4

(FACTOR1 -, FACTOR2 +)

RURAL

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The vast majority (81.5 percent) of New England counties' scores fall into QUAD2, 17 percent fall in QUAD3, and 1.5 percent are located in QUAD1. The 1.5 percent in QUAD1 represents a single occurrence, Suffolk County (Boston), Massachusetts. 83 percent of New England counties are more urban than the mean county, and 98.5 percent are ethnically more foreign born and Roman Catholic than the mean county. The scores of the Middle Atlantic states (New York, New Jersey, Pennsylvania, and Delaware) are similarly distributed; 77 percent of the Middle Atlantic counties are in QUAD2.

The Old Northwest Territory states (Ohio, Indiana, Michigan, Illinois, and Wisconsin) yield scores that are much more evenly distributed between QUAD2 and QUAD3 than either of the preceding regions. Each of these quadrants holds approximately 40 percent of the region's counties. 64 (15 percent) of the Old Northwest Territory counties score in QUAD4. Of these 64 counties 19 are in Illinois, 28 in Indiana, 15 in Ohio, and two are in Michigan. QUAD4 counties are concentrated in the southern portions of these states. The eight Northwest Territory counties in QUAD1 represent the region's urban centers.

The plains states (Kansas, Nebraska, South Dakota, North Dakota, Minnesota, and Iowa) are predominantly rural and ethnically foreign born and Roman Catholic; 80.7 percent of the region's 441 counties are in QUAD3. Of the Plains counties, the 15.7 percent that were urban all placed in QUAD2. The scores of the few Plains counties not already accounted for, 16 (3.6 percent), place them in QUAD4.

The Southern states (Virginia, North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, Louisiana, Arkansas, and
Texas) and the Border States (Maryland, West Virginia, Kentucky, Tennessee, and Missouri) compose the positive scope of FACTOR2. 94.9 percent of all QUAD1 counties lie in these two regions, as do 89.8 percent of all QUAD4 counties. Further, 90.8 percent of Southern counties place in either QUAD1 or QUAD4, as do 84.5 percent of the Border counties. The majority of Southern counties (55.95 percent) are classified into QUAD4. Only 34.9 percent of Southern counties and 10.1 percent of Border counties fall in QUAD1. Of the 79 Southern counties in QUAD2 and QUAD3, 56 from Texas (15 in QUAD2 and 41 in QUAD3). This distribution is expected in the city poor South because QUAD1 contains the cities and QUAD4 contains the rural counties. Texas is identified as a marginal element of the region.

The two Western regions are very similar. The Mountain Western states (Montana, Idaho, Wyoming, Colorado, Utah, and Nevada) are overwhelmingly foreign born, Roman Catholic, and strongly rural. 76 percent of the Western counties lie in QUAD3, as one would expect. Pacific West Coast (Washington, Oregon, and California) counties are much more evenly distributed between QUAD2 and QUAD3, 42 percent and 58 percent of the regions counties are in each, respectively. Only one Mountain Western county is located in QUAD4, and there are no Pacific or Western counties in QUAD1.

The scores of counties in the socio-economic dimensions reflect the structure of the similarites that exist between the socio-economic and spatial dimensions. Any attempt to model in one set of dimensions without the influence of the other cannot be wholly successful due to the inherent covariance. The degree of covariance demonstrates that similarity and propinquity are measures of much the same thing, and
that, in this case, neighbors tend to very much alike. One might suggest that the relationship between the spatial variables, the urban variable and the ethnic-racial-religious variable, have defined culture areas within the United States. The organization and distribution of counties in the social dimensions is, however, substantially transformed from that of the spatial dimensions despite the strength of the relationship between the dimensions.

5.3 Assignment of Socio-economic Predictive Generations to Specific Observed Years

As with the spatial dimensions, it is necessary to assign the results of each scanner generation from the omega pattern to a year from the observed pattern. The same criteria are used to assign predictions to observations; the reduction of the summed difference between observed and predicted percent dry. The years that minimize this difference are: 1884, 1887, 1892, 1908, 1909, 1910, 1916, 1917, and 1918 (Table 5.4). The correlation between the predicted dry percentages and the observed dry percentages is .99279. This series is not quite as accurately matched as the spatial series. Despite the minor reduction in conformity, predictive accuracy is not impaired.

The model behaves in a manner identical to the model designed to simulate diffusion in the spatial dimensions. Accuracy is highest closest to 1876, the base year. The model was 92.1 percent accurate in 1884, the first predicted for year. The rate of accuracy experiences its trough at 47.6 percent in 1908. As the movement saturates, the nation accuracy again climbs reaching 67.5 percent in 1918. Overall, the model when implemented on the socio-economic dimensions is
<table>
<thead>
<tr>
<th>Year</th>
<th>Percent Observed Dry</th>
<th>Percent Predicted Dry</th>
<th>Percent Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1876</td>
<td>9.8624</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1877</td>
<td>11.2768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1878</td>
<td>11.0474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1879</td>
<td>11.3150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1880</td>
<td>11.5826</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1881</td>
<td>15.5963</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1882</td>
<td>15.8640</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1883</td>
<td>16.5520</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1884</td>
<td>17.5841</td>
<td>17.508</td>
<td>92.106</td>
</tr>
<tr>
<td>1885</td>
<td>18.4633</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1886</td>
<td>19.5366</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1887</td>
<td>24.8471</td>
<td>24.809</td>
<td>73.583</td>
</tr>
<tr>
<td>1888</td>
<td>26.1850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1889</td>
<td>27.4464</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1890</td>
<td>27.5229</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1891</td>
<td>28.0581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1892</td>
<td>28.2492</td>
<td>31.384</td>
<td>63.910</td>
</tr>
<tr>
<td>1893</td>
<td>24.0382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1894</td>
<td>25.0382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1895</td>
<td>23.3180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1896</td>
<td>23.0505</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1897</td>
<td>21.4450</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1898</td>
<td>21.4067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1899</td>
<td>21.7900</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1900</td>
<td>22.7064</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1901</td>
<td>22.6300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1902</td>
<td>24.6177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1903</td>
<td>25.2676</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1904</td>
<td>36.4297</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1905</td>
<td>36.3532</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>36.2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1907</td>
<td>38.7615</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1908</td>
<td>45.8716</td>
<td>40.635</td>
<td>47.628</td>
</tr>
<tr>
<td>1909</td>
<td>55.0077</td>
<td>51.070</td>
<td>48.778</td>
</tr>
<tr>
<td>1910</td>
<td>59.4037</td>
<td>60.130</td>
<td>51.377</td>
</tr>
<tr>
<td>1911</td>
<td>64.1055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1912</td>
<td>62.1177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1913</td>
<td>61.2768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1914</td>
<td>58.7156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>61.6208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1916</td>
<td>67.2232</td>
<td>68.387</td>
<td>53.940</td>
</tr>
<tr>
<td>1917</td>
<td>75.2294</td>
<td>76.109</td>
<td>60.803</td>
</tr>
<tr>
<td>1918</td>
<td>79.8165</td>
<td>82.148</td>
<td>67.533</td>
</tr>
<tr>
<td>1919</td>
<td>100.0000</td>
<td>100.000</td>
<td></td>
</tr>
</tbody>
</table>

\[ r(\text{percent observed dry, percent predicted dry}) = .99279 \]

sum 559.656
mean 62.184

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
62.1 percent accurate, eight percent less accurate than the spatial dimensions.

Unlike the predictions based on the spatial dimensions, however, the rate of growth did not slow towards the end of the simulation. Dry counties in 1876 are much more scattered in the socio-economic dimensions compared to the spatial dimensions (compare Fig. 5.1 and Fig. 4.1). The greater dispersal of counties observed dry promotes more rapid predicted growth in the early generations even though the same number of counties are initially dry (less those excluded when the study population dropped from 3030 to 2499 counties) because proportionally more conversion signals are sent to wet counties.

In Iowa, a dry state in 1876, for example, only the counties on the borders sent effective conversion signals in the spatial dimensions, the interior counties merely sent conversion signals to each other. In the socio-economic dimensions, however, the counties of Iowa are interspersed with wet counties from all over the northeastern United States. When Iowan counties transmit conversion signals in the social model each has a higher probability of contacting a wet county than the same county does in the spatial model. In one way this is an advantage because the scale problem noted when a state adopts is partially counteracted. Contiguous blocks no longer adopt, but counties that would not have adopted on their own are still unaccountable.

This interpretation of events is supported by noting two facts. The first scanner generation of the spatial dimensions matched with the year 1878, a growth of 1.19 percent for the first scan, but the socio-economic dimensions first scanner generation produced so many new dry counties that their first matching year does not occur until
Figure 5.1
1884, a growth of 7.65 percent. Second, the slow growth of the spatial model permitted twelve scanner generations to be matched, but the social model only permitted nine. Further, there were several unmatchable generations in the later years of the spatial model. The breakup of dry clusters through reorganization permits continued predictions of rapid spread, even to the termination of the simulation model when eighty percent of the nation's counties were predicted dry.

A second reason for the slowing of growth in the spatial dimensions and not in the socio-economic is the differential role, treatment, and position of the Western counties in the distributions. In the spatial dimensions, the Western and Pacific counties are probabilistically isolated. This is due to the number of intervening opportunities for county conversion between Western counties and the nearest dry county to them in 1876 and, potentially, to the radical shift in density of counties west of the Rockies. The model using the spatial dimensions predicts saturation east of the Rockies, but predicted growth did not pass beyond the mountain barrier before the eighty percent cutoff was reached. In the social dimensions, on the other hand, the Western counties are more centrally located in the distribution and the probability of contact is not shifted away from them. The characteristic asymptotic relationship between the cumulative frequency curve and the line of total adoption that develops as the simulation approaches saturation is not observed in either social or spatial models. This is because simulated national prohibition as employed here is below the point on the curve where expected flattening occurs.

While the number of observed and predicted dry counties are
closely matched for the entire nation, there are substantial differences between the observed and predicted numbers of dry counties in three of the four quadrants (Table 5.5). QUAD1 and QUAD4 counties tend to adopt sooner, and in greater numbers, than predicted. The mean observed percent dry for QUAD4 is 68.7 percent and the mean predicted percent dry is 47.91 percent, a difference of 20.79 percent. For QUAD1 the same figures, 57.95 percent mean observed dry and 33.25 percent mean predicted dry, results in a difference of 33.25 percent. Growth is more closely monitored and predicted in QUAD4, where the mean accuracy is 64.86 percent, than in QUAD1 where it is only 58.52 percent. The high rate of overprediction in these quadrants is verified by the type I error to total error ratio which for each quadrant has a mean value below 0.25 (Less than one error in four were type I).

Among QUAD2 counties predicted rates of acceptance deviate more from the observed numbers than for any other quadrant. A gap of 30.37 percentage points exists between the two statistics. Further, for QUAD2 mean accuracy is the lowest of any quadrant, 45.51 percent, and QUAD2's type I to total error ratio is 0.877. The low accuracy rate depicts extensive failure and the type I to total ratio indicates the direction of that failure, overprediction.

QUAD3 stands apart from the other quadrants because of the close similarity between the mean observed and predicted percent of dry counties, only 4.8 percentage points separate them. Accuracy for QUAD3 is just over that for QUAD4, 64.96 percent, but 56.23 percent of all QUAD3 errors are type I indicating that the two are otherwise dissimilar.

Matching predictions to observations for the entire population of
Table 5.5: Accuracy of the Socio-Economic Dimension, by Quadrant

<table>
<thead>
<tr>
<th>Year</th>
<th>Observed</th>
<th>Predicted</th>
<th>Correct</th>
<th>Errors of Type I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
</tr>
<tr>
<td></td>
<td>Dry</td>
<td>Predicted Dry</td>
<td>Correct</td>
<td>of Errors</td>
</tr>
<tr>
<td>1884</td>
<td>6.91</td>
<td>5.11</td>
<td>92.19</td>
<td>38.46</td>
</tr>
<tr>
<td>1887</td>
<td>19.29</td>
<td>9.79</td>
<td>78.64</td>
<td>27.78</td>
</tr>
<tr>
<td>1892</td>
<td>20.12</td>
<td>15.87</td>
<td>73.08</td>
<td>41.76</td>
</tr>
<tr>
<td>1908</td>
<td>61.47</td>
<td>22.65</td>
<td>40.59</td>
<td>17.33</td>
</tr>
<tr>
<td>1909</td>
<td>74.71</td>
<td>30.83</td>
<td>37.35</td>
<td>15.02</td>
</tr>
<tr>
<td>1910</td>
<td>80.06</td>
<td>40.76</td>
<td>42.65</td>
<td>15.82</td>
</tr>
<tr>
<td>1916</td>
<td>84.16</td>
<td>48.68</td>
<td>44.28</td>
<td>17.99</td>
</tr>
<tr>
<td>1917</td>
<td>85.04</td>
<td>58.36</td>
<td>53.37</td>
<td>21.38</td>
</tr>
<tr>
<td>1918</td>
<td>89.74</td>
<td>67.16</td>
<td>64.52</td>
<td>18.18</td>
</tr>
<tr>
<td>Mean</td>
<td>57.95</td>
<td>33.25</td>
<td>58.52</td>
<td>23.75</td>
</tr>
</tbody>
</table>

Observed Mean - Predicted Mean= 24.7

<table>
<thead>
<tr>
<th>Year</th>
<th>Observed</th>
<th>Predicted</th>
<th>Correct</th>
<th>Errors of Type I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
</tr>
<tr>
<td></td>
<td>Dry</td>
<td>Predicted Dry</td>
<td>Correct</td>
<td>of Errors</td>
</tr>
<tr>
<td>1884</td>
<td>20.83</td>
<td>31.04</td>
<td>85.07</td>
<td>84.21</td>
</tr>
<tr>
<td>1887</td>
<td>22.52</td>
<td>40.58</td>
<td>74.95</td>
<td>86.05</td>
</tr>
<tr>
<td>1892</td>
<td>22.63</td>
<td>48.16</td>
<td>67.12</td>
<td>88.82</td>
</tr>
<tr>
<td>1908</td>
<td>15.30</td>
<td>57.74</td>
<td>47.23</td>
<td>90.22</td>
</tr>
<tr>
<td>1909</td>
<td>19.69</td>
<td>67.69</td>
<td>43.21</td>
<td>92.26</td>
</tr>
<tr>
<td>1910</td>
<td>21.80</td>
<td>73.81</td>
<td>41.30</td>
<td>94.14</td>
</tr>
<tr>
<td>1916</td>
<td>27.53</td>
<td>79.16</td>
<td>37.67</td>
<td>91.41</td>
</tr>
<tr>
<td>1917</td>
<td>38.05</td>
<td>85.34</td>
<td>41.49</td>
<td>90.20</td>
</tr>
<tr>
<td>1918</td>
<td>47.23</td>
<td>88.34</td>
<td>45.51</td>
<td>87.72</td>
</tr>
<tr>
<td>Mean</td>
<td>26.73</td>
<td>57.10</td>
<td>53.73</td>
<td>89.45</td>
</tr>
</tbody>
</table>

Observed Mean - Predicted Mean= -30.37

1The number of observations for years prior to 1896 is less than the indicated number because not all counties used in the analysis were formed prior to that date.
Table 5.5: Continued

QUAD3  \( n = 780 \)

<table>
<thead>
<tr>
<th>Year</th>
<th>Percent Observed Dry</th>
<th>Percent Predicted Dry</th>
<th>Percent Correct</th>
<th>Percent of Errors Type I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1884</td>
<td>24.48</td>
<td>23.24</td>
<td>84.09</td>
<td>46.09</td>
</tr>
<tr>
<td>1887</td>
<td>25.47</td>
<td>29.92</td>
<td>77.49</td>
<td>59.88</td>
</tr>
<tr>
<td>1892</td>
<td>34.95</td>
<td>34.29</td>
<td>66.62</td>
<td>49.02</td>
</tr>
<tr>
<td>1908</td>
<td>31.41</td>
<td>43.46</td>
<td>60.00</td>
<td>65.06</td>
</tr>
<tr>
<td>1909</td>
<td>38.46</td>
<td>52.31</td>
<td>57.18</td>
<td>66.17</td>
</tr>
<tr>
<td>1910</td>
<td>44.87</td>
<td>60.26</td>
<td>56.40</td>
<td>67.65</td>
</tr>
<tr>
<td>1916</td>
<td>60.06</td>
<td>67.18</td>
<td>56.03</td>
<td>57.43</td>
</tr>
<tr>
<td>1917</td>
<td>78.40</td>
<td>73.33</td>
<td>60.77</td>
<td>46.08</td>
</tr>
<tr>
<td>1918</td>
<td>81.73</td>
<td>79.10</td>
<td>66.03</td>
<td>48.69</td>
</tr>
</tbody>
</table>

Mean 46.65 51.45 64.96 56.23

Observed Mean - Predicted Mean = 4.80

QUAD4  \( n = 855 \)

<table>
<thead>
<tr>
<th>Year</th>
<th>Percent Observed Dry</th>
<th>Percent Predicted Dry</th>
<th>Percent Correct</th>
<th>Percent of Errors Type I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1884</td>
<td>17.38</td>
<td>11.23</td>
<td>82.98</td>
<td>31.94</td>
</tr>
<tr>
<td>1887</td>
<td>31.22</td>
<td>17.72</td>
<td>67.02</td>
<td>29.54</td>
</tr>
<tr>
<td>1892</td>
<td>31.82</td>
<td>24.82</td>
<td>61.94</td>
<td>41.23</td>
</tr>
<tr>
<td>1908</td>
<td>74.15</td>
<td>34.85</td>
<td>38.48</td>
<td>18.06</td>
</tr>
<tr>
<td>1909</td>
<td>86.67</td>
<td>47.37</td>
<td>48.54</td>
<td>11.82</td>
</tr>
<tr>
<td>1910</td>
<td>90.41</td>
<td>58.71</td>
<td>56.84</td>
<td>13.28</td>
</tr>
<tr>
<td>1916</td>
<td>94.04</td>
<td>70.06</td>
<td>66.43</td>
<td>14.29</td>
</tr>
<tr>
<td>1917</td>
<td>95.09</td>
<td>79.65</td>
<td>76.84</td>
<td>17.65</td>
</tr>
<tr>
<td>1918</td>
<td>97.54</td>
<td>86.90</td>
<td>84.67</td>
<td>15.79</td>
</tr>
</tbody>
</table>

Mean 68.70 47.91 64.86 21.51

Observed Mean - Predicted Mean = 20.79

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
counties does not insure that any sample of counties will match equally well. The expectation is, in fact, the opposite because the proportion dry in the population is, by definition, the mean proportion of all possible samples. The difference between the observed and predicted incidence of dry counties in any sample can identify trends in both the entire population and the subgroup that might otherwise have gone undetected.

The subsetting of the counties included in the analysis have yielded several pieces of information. There is a strong tendency to overpredict adoption by counties that are urban with higher than average proportions of foreign born and Roman Catholic residents. Conversely, there is an equally strong trend to underpredict growth in counties with a higher than average proportion of black residents. Further, adoption by rural counties is more closely and accurately predicted than by urban counties. Predictions match much more closely for rural counties with a high proportion of foreign born and Roman Catholic residents than those with equal proportions of black residents, but they are equally well predicted.

5.4 Analysis of the Distribution of Predictive Errors

The distribution of errors generated by the social model is not random. The arrangement of errors is the result of systemic model failure from several sources. An understanding of conditions that result in errors leads directly to a comprehension of error sources and of the shortcomings of the constituent elements of the dimensions. Once points and areas of predictive errors are identified, explanations for their occurrence, and perhaps their elimination, may be
undertaken.

The distribution of wet and dry counties in 1876 forms the basis for the model’s predictions in the social model, just as it did for the spatial model. In all cases when the number and percent increase in dry counties is given it is the net increase. In many instances, because of the nature of prohibition’s adoption, a county may be prohibition territory one year and wet the next. There was much more adoptive and recidivous activity than can be expressed in the summary statistics presented here. Some problems with recidivism are avoided by referring to statistics based on permanent adoption. Permanent adopters are those counties that after adoption remain dry throughout the remainder of the study period.

5.41 Predictions for 1884

Between 1876 and 1884 over 200 counties adopted prohibition, bringing the total number of dry counties to 465. Adoption by the state of Kansas in 1881 accounted for 95 of these newly dry counties. All 73 of the new adopters in QUAD3 are Kansan counties. Kansan adoption also accounts for twelve of the thirteen new adopters in QUAD2 and nine adopters in QUAD4. Of the remaining 91 QUAD4 adopters the majority are Southern. 82 of the 91 QUAD4 adopters were permanent adopters. Of these 26 are Georgian, 11 Mississippian, 31 Kentuckian, and 7 Tennesseen. Both Georgia and Kentucky were sites of active, religiously oriented, change agents at this time. All newly dry QUAD1 counties are from the South, eleven of which are permanent adopters from Georgia. Two of these are the ante-bellum local option counties of Liberty and Camden, both of which adopted in 1877.
As expected, the fewest predictive errors occur for 1884 because it is the first year for which predictions are cast. Predictive errors are almost equally split between type I (counties that were wet, but predicted dry) and type II errors (Table 5.6). Type I errors are adjacent to correctly predicted dry counties, just as they were in the predictions of the spatial model, except that this time the counties are adjacent because they are similar. 173, 6.9 percent, of all counties are type I errors in 1884's predictions. There are 64 type I errors in QUAD2, 53 in QUAD3, and 46 in QUAD4. There are 188 type II errors, representing 7.5 percent of all counties. Type II errors are concentrated in QUAD3 and QUAD4. In QUAD4 the number of dry counties is underpredicted, but in QUAD3 dry counties are mispredicted.

A breakdown of the distribution of type I and type II errors by quadrant indicates a trend in their locations within the social dimensions. By 1884, 6.9 percent (25) of all QUAD1 counties had adopted. The model predicts that only 5.1 percent (17) are adopters. Of those 17, however, only seven were really dry (one dry county in 1876 had returned to the wet status), the other ten are type I errors. Ten type I errors are located in QUAD1 where counties are more urban and have higher concentrations of black residents than the mean county. Because there were only eight actual dry counties in QUAD1 in 1876 the other two of the ten are generated by dry counties on the QUAD1 border. This leaves 18 type II errors.

106 QUAD2 counties were observed dry in 1884, representing 20.8 percent of the 509 counties in QUAD2. Many more QUAD2 counties are predicted dry, 158 in all. In 94 cases the predicted dry status is correct, but 93 of these are given because they were dry in 1876. The
### Table 5.6: Accuracy of the Socio-Economic Dimensions in 1884, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry*</td>
</tr>
<tr>
<td>dry</td>
<td>wet**</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>31.04</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>20.83</td>
</tr>
<tr>
<td>number observed dry</td>
<td>106</td>
</tr>
<tr>
<td>percent correct</td>
<td>85.07</td>
</tr>
<tr>
<td>n=509</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry</td>
</tr>
<tr>
<td>dry</td>
<td>wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>23.24</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>24.48</td>
</tr>
<tr>
<td>number observed dry</td>
<td>177</td>
</tr>
<tr>
<td>percent correct</td>
<td>84.09</td>
</tr>
<tr>
<td>n=723</td>
<td></td>
</tr>
</tbody>
</table>

* Type I error
** Type II error

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
model predicts that 65 QUAD2 counties adopted between 1876 and 1884. In reality, only 13 adopted and the model only correctly identifies one of them.

The distribution of errors in QUAD3 is split very evenly between the two types of errors, 46.09 percent of the 115 errors are type I. 115 counties are correctly predicted dry, up eleven from 1876. In QUAD3 the model is more accurate in selecting counties that adopted than in QUAD2 despite the fact that QUAD2 is more accurately predicted. In QUAD3 73 counties adopted between 1876 and 1884 making the total number of actual dry counties in QUAD3 177, or 24.5 percent of all QUAD3 counties. The model, however, predicts that only 23.2 percent, 168, are dry, up 64 counties from 1876. The model is wrong about which counties would adopt 53 times. These 53 mistakes are type I errors. The 62 counties that adopted and are not identified by the model are type II errors. Eleven hits out of 64 predictions is a fairly good average given the great potential for error, and it is certainly much better than the predictions for QUAD2.

In QUAD4 more than twice as many type II errors occur as type I errors, 31.97 percent are type I. 147, 17.4 percent, of QUAD4 counties were dry in 1884, but only 95 are predicted dry. Of the 147 observed dry counties 49, up two from 1876, are predicted dry. The remaining 98 dry counties are predicted wet. There are 46 type I errors in QUAD4 and 98 type II errors.

What is the cause of the more rapid than predicted growth among counties with higher than average black populations? And of the slower than expected growth among QUAD2 counties? During the 1850s many QUAD2 counties adopted prohibition. Those that were still dry in
1876 were the 'hangers on'. They, unlike so many of their com-patriots, lacked the impetus to repeal, or, perhaps, these counties and states were content with prohibition, at that time. Regardless, the model was not designed to accommodate the fact that growth was not present among any subset of the population. Growth in QUAD1 and QUAD4 is more rapid than expected because religious leaders in these types of counties embraced the movement during the late 1870s and early 1880s. Adoption by QUAD1 and QUAD4 counties results from their actions and influences. The model is unable to account for this relocation event and a great number of errors result.

The structure of the model also has an influence. Each predicted years growth is matched to an observed year based on the number of dry counties in each. When 203 new counties went dry between 1876 and 1884, and the model expected 65 in QUAD2 and there were only 13; QUAD1 and QUAD4 had to take up the slack. Overprediction in one subgroup must be compensated for by underprediction in another.

5.42 Predictions for 1887

The distribution of dry counties in 1887 was an outgrowth of the trends first identified in 1884 (Table 5.7). In QUAD1 growth was accelerating. Forty additional QUAD1 counties adopted during the three years 1884 to 1887 bringing the proportion of dry counties in QUAD1 to 19.3 percent. Fifteen of the 26 permanent adopters were from Georgia, and four from Mississippi, only Warren County, Kentucky was from outside the South.

Growth in QUAD2, on the other hand, was very slow. Only ten QUAD2 counties adopted between 1884 and 1887, bringing the total
Table 5.7: Accuracy of the Socio-Economic Dimensions in 1887, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs pred freq percent</td>
<td>obs pred freq percent</td>
</tr>
<tr>
<td>wet wet 288 55.92</td>
<td>wet wet 252 74.78</td>
</tr>
<tr>
<td>dry dry 98 19.03</td>
<td>dry dry 13 3.86</td>
</tr>
<tr>
<td>wet dry* 111 21.55</td>
<td>wet dry 20 5.94</td>
</tr>
<tr>
<td>dry wet** 18 3.49</td>
<td>dry wet 52 15.43</td>
</tr>
<tr>
<td>percent predicted dry 40.58</td>
<td>percent predicted dry 9.79</td>
</tr>
<tr>
<td>percent observed dry 22.52</td>
<td>percent observed dry 19.29</td>
</tr>
<tr>
<td>number observed dry 116</td>
<td>number observed dry 65</td>
</tr>
<tr>
<td>percent correct 74.95</td>
<td>percent correct 78.64</td>
</tr>
<tr>
<td>n=515</td>
<td>n=337</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs pred freq percent</td>
<td>obs pred freq percent</td>
</tr>
<tr>
<td>wet wet 453 61.05</td>
<td>wet wet 503 59.04</td>
</tr>
<tr>
<td>dry dry 122 16.44</td>
<td>dry dry 68 7.98</td>
</tr>
<tr>
<td>wet dry 100 13.48</td>
<td>wet dry 83 9.74</td>
</tr>
<tr>
<td>dry wet 67 9.03</td>
<td>dry wet 198 23.24</td>
</tr>
<tr>
<td>percent predicted dry 29.92</td>
<td>percent predicted dry 17.72</td>
</tr>
<tr>
<td>percent observed dry 25.47</td>
<td>percent observed dry 31.22</td>
</tr>
<tr>
<td>number observed dry 189</td>
<td>number observed dry 266</td>
</tr>
<tr>
<td>percent correct 77.49</td>
<td>percent correct 67.02</td>
</tr>
<tr>
<td>n=742</td>
<td>n=852</td>
</tr>
</tbody>
</table>

* Type I error
** Type II error
number dry in QUAD2 to 116. In QUAD3 growth was also slow, only 12 additional counties adopted. Three of the newly dry counties in QUAD3 resulted from the formation of new counties in Kansas. These three counties represent the only permanent adopters outside the South and Border regions during this period. Five of the remaining nine QUAD3 adopters were from West Virginia.

Growth in QUAD4 was very rapid. During these three years 115 counties adopted. 31.2 percent of all QUAD4 counties were dry in 1887 up 13.8 percent from 1884. All QUAD4 adopters are from either the South or Border regions, and Georgia, Kentucky, and West Virginia each contributed over 20 new adopters. Many counties in QUAD1 and QUAD4 were under pressure to adopt from religious leaders. It is hypothesized that success in proselytizing activity by religious leaders and the amount of such activity among QUAD1 and QUAD4 counties varied with level of urbanization. In rural counties the growth rate was higher than in urban counties. The close-knit relationships and social control that existed in rural places perhaps made it easy to shame people into doing the moral, god [i.e. clergy] approved thing. In urban places, however, with less social control, and more bars, distillaries, and breweries the religious leaders had more difficulty. Perhaps the urban clergy were less likely to attempt persuading their congregations to abandon demon rum because they knew it would be an unpopular topic. were also The small difference in the growth rates of QUAD1 and QUAD4 is attributed to the variable that distinguishes the two groups, level of urbanization.

In 1885, there were 16 permanent adopters, ten from QUAD4, four from QUAD1, and two of the three Kansan counties from QUAD3. Data
from the Census of 1890 shows that the mean percent black for those counties was 35.39 percent. In 1886, when there were 17 permanent adopters, 13 from QUAD4 and 4 from QUAD1, the mean percent black in 1890 was 34.96 percent. There were many more adopters in 1887, 97 in all, 67 from QUAD4, six from QUAD3, five from QUAD2, and 19 from QUAD1. The increased representation of QUAD2 and QUAD3 had a strong influence on the mean percentage black from the 1890 census which was 18.84 percent. The other ethnic census variables were apparently unaffected by this change in the mix of adopters.

The predictive tendencies first noted for 1884 continue. In QUAD1 the number of predicted dry counties increase by 16, much less than the 40 observed. The result is that 15.4 percent of QUAD1 counties are categorized as type II errors, but only 5.9 percent are type I. QUAD1 predictions are, therefore, 78.6 percent accurate. Type I errors account for 111 QUAD2 and 100 QUAD3 counties. Type IIIs represent a mere 18 QUAD2 and 67 QUAD3 counties. Overprediction in QUAD3 is very small, but in QUAD2 it is becoming a very serious source of error; almost as serious as the underprediction in QUAD4. In QUAD4 there are 198 type II errors and only 83 type I. The relocation of prohibition to QUAD4 from QUAD2 after the Civil War thanks to the activity of Southern preachers has begun to have a severe negative effect on the predictive capability of the social model.

5.43 Predictions for 1892

The pattern of adoption in 1892 was, because of the Dakotas, very different from that observed in 1887. Adoption by the Dakotas represented an increase of 75 dry counties in QUAD3 and seven in QUAD2.
Only three other QUAD3 counties adopted between 1887 and 1892 (Table 5.8). In QUAD2 even though seven new counties adopted six that were previously dry repealed their dry status for a net gain of only one county. There were only three additional dry counties in each of the remaining quadrants during this five year period.

The vastly different growth patterns during these five years is noticeable in the values of the ethnicity variables. The mean percent of county residents foreign born in 1890 is 36.25 percent for counties that adopted permanently in 1889, the year the Dakotas adopted. The same statistic for 1887 is 1.32 percent. The percent Roman Catholic 1900 statistic is similarly affected. 7.93 percent of the county residents for 1889 adopters were Roman Catholics but only 0.25 percent of county residents in 1887 adopters were Roman Catholic. The percent Negro in 1890 is, naturally, affected oppositely dropping to 4.15 percent among adopters in 1889. Because ethnic, racial, and religious variables are different and not indicators of urbanness the shift in adopter location in the dimensions is primarily a lateral one. The focus of adoption activity shifting from QUAD4 to QUAD3.

The observed growth in QUAD1 continues to vastly outstrip the predicted rate of growth. QUAD1 was 19.3 percent dry in 1887 and 20.1 percent dry in 1892, but the model predicts that QUAD1 is only 9.8 percent and 15.9 percent dry for each of the years. In both 1887 and 1892, type II errors in QUAD1 remain practically constant at 15.5 percent of all QUAD1 counties. Type I errors climb rapidly, however, and their increase results in a reduction of accuracy within the quadrant of 5.5 percent between the 1887 and 1892 predictions.

The situation in QUAD4 is almost identical to that of QUAD1.
Table 5.8: Accuracy of the Socio-Economic Dimensions in 1892, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obs</strong></td>
<td><strong>pred</strong></td>
</tr>
<tr>
<td>wet wet</td>
<td>249</td>
</tr>
<tr>
<td>dry dry</td>
<td>98</td>
</tr>
<tr>
<td>wet dry*</td>
<td>151</td>
</tr>
<tr>
<td>dry wet**</td>
<td>19</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>48.16</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>22.63</td>
</tr>
<tr>
<td>number observed dry</td>
<td>117</td>
</tr>
<tr>
<td>percent correct</td>
<td>67.12</td>
</tr>
<tr>
<td>n=517</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obs</strong></td>
<td><strong>pred</strong></td>
</tr>
<tr>
<td>wet wet</td>
<td>372</td>
</tr>
<tr>
<td>dry dry</td>
<td>137</td>
</tr>
<tr>
<td>wet dry</td>
<td>125</td>
</tr>
<tr>
<td>dry wet</td>
<td>130</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>34.29</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>34.95</td>
</tr>
<tr>
<td>number observed dry</td>
<td>267</td>
</tr>
<tr>
<td>percent correct</td>
<td>66.62</td>
</tr>
<tr>
<td>n=764</td>
<td></td>
</tr>
</tbody>
</table>

* Type I error
** Type II error
reflecting a general trend for the model to underestimate dry territory in the South and Border regions. Between 1887 and 1892 the incidence of dry counties in QUAD4 remains an almost constant 31.5 percent. The increase in the number of predicted dry counties between 1887 and 1892, which is substantial, does little to reduce the number of type II errors. There were only seven fewer type II errors in 1892 than in 1887. The remaining model converted counties are represented as type I errors. In both QUAD1 and QUAD4 the observed loss of accuracy during the period 1887 to 1892 is the result of the model's failure to select the correct counties to convert coupled with an underprediction of the number of dry counties. QUAD2 displays the reverse situation in predicted and observed rates of spread. Type I errors occur in QUAD2 at seven times the rate of type II errors. Owing to the many counties originally dry, the predicted incidence of dry counties climbs from 40.5 percent in 1887 to 48.2 percent in 1892, but the observed incidence increases only 0.1 percent during the same period. The percentage of type I errors in QUAD2 increases by 7.7 percent from 1887 to 1892, the same amount that dry predictions increased by.

QUAD3 is set apart from the other groups of counties because of the consistent near equality observed in the observed and predicted incidence dry. In 1884, QUAD3 was 24.5 percent dry, and it is predicted to be 23.2 percent dry, a difference of only 1.2 percent. In 1892 the difference between observed and predicted percent dry is 0.7 percent. Despite this high covariance, there is a strong pattern of errors. In 1892, type II errors slightly surpass the number of type I errors. Predictions for QUAD3 counties in 1892 are 66.6 percent
accurate, a drop of nearly 11 percent from the 1887 predictions. Because the errors are evenly distributed between the two types, and continue to be evenly distributed throughout the study period, the predictive trend is assumed to be more in line with the observed trend in QUAD3 than in any other group of the distribution.

The concentration of growth in QUAD3 during this period has permitted predictions to catch up in QUAD1 and QUAD4. Despite the fact that the figures are more in line with each other, accuracy decreases because the model misidentifies adopters more frequently than it correctly identifies them. Simultaneously, the territorial gains in QUAD3 caused an increased divergence between the incidence of predicted and observed dry counties in QUAD2.

5.44 Predictions for 1908

1908 is the first year after 1892 that predictions are cast for. The previously cited decline in the number of dry counties and the inability of the model to predict decline are responsible for the 16 year gap in predictions. All of the factors that were discussed in the 1904 analysis of the spatial model apply here. Their impacts, however, are perceived differently (Table 5.9).

The reduction of dry territory that occurred as a result of repeal by New Hampshire, Vermont, Ohio, Iowa, and South Dakota between 1892 and 1903 impacted counties with higher than average proportions of foreign born residents greatly. 83 counties from these five states are located in QUAD2, and 146 are in QUAD3. Only 21 counties, all from Ohio and Iowa, are in QUAD4, and two Ohio counties, Greene and Ross, are located in QUAD1. Some counties from these five states
Table 5.9: Accuracy of the Socio-Economic Dimensions in 1908, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obs</strong></td>
<td><strong>pred</strong></td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry</td>
</tr>
<tr>
<td>dry</td>
<td>wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>57.74</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>15.30</td>
</tr>
<tr>
<td>number observed dry</td>
<td>80</td>
</tr>
<tr>
<td>percent correct</td>
<td>47.23</td>
</tr>
<tr>
<td>n=523</td>
<td>n=341</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>obs</strong></td>
<td><strong>pred</strong></td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry</td>
</tr>
<tr>
<td>dry</td>
<td>wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>43.46</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>31.41</td>
</tr>
<tr>
<td>number observed dry</td>
<td>249</td>
</tr>
<tr>
<td>percent correct</td>
<td>60.00</td>
</tr>
<tr>
<td>n=780</td>
<td>n=855</td>
</tr>
</tbody>
</table>

* Type I error
** Type II error

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
retained local option, but not many. 35 QUAD3 and four QUAD4 counties from Iowa retained dry status after repeal and into national prohibition, for example, but Iowa is exceptional. In QUAD2 between 1892 and 1908, there was a net loss of 37 dry counties, but QUAD3 had a net loss of only 18 dry counties. A large amount of status shifting is evident in QUAD2 and QUAD3. Territorial loss is difficult to identify because much of the lost ground was recovered by adoptions after 1902.

Between 1892 and 1908 growth in QUAD1 and QUAD4 was phenomenal. 365 QUAD4 and 141 QUAD1 counties adopted during this period. Growth was not constant over the 16 year period, obviously. It was very slow in the mid 1890s, but by 1902 the take off point had been reached in these two quadrants. All 19 counties that adopted between 1893 and 1897 and stayed dry through 1920 are from either QUAD1 or QUAD4. Every county that adopted permanently between 1893 and 1897 was over 61 percent black, based on data from the Census of 1890. Eight of these are QUAD1 counties from Mississippi with an average ethnicity score of 2.08. Five are QUAD4 counties from Virginia with an average ethnicity score of 0.605. The overall ethnicity mean for these 19 counties is 1.572. There are many counties within this range, but among this highly selective group the relationship between ethnicity and urbanness scores is almost linear; the more urban a county the higher its ethnicity score.

Oklahoma's adoption in 1907 acted as a catalyst, sparking adoption throughout the nation. In 1908 Georgia, which was over 80 percent dry in 1907, adopted. Half, ten, of the Georgian counties that were forced to adopt in 1908 were from QUAD1, but the mix of Georgian counties is 59 QUAD1 and 71 QUAD4. None of Georgia's QUAD1 counties
are terribly urban. Fulton County, which adopted in 1905, only scores 2.078 on the urban scale and 1.044 on the ethnicity scale, is Georgia's most urban county. Of those QUAD1 counties that adopted in 1908, the mean level of urbanness was 0.403, standard deviation 0.307, and a mean ethnicity of 1.636, standard deviation 0.377. Laggard QUAD4 counties had a mean urbanness of -0.255, standard deviation 0.205, and mean ethnicity of 1.104, standard deviation 0.524. While these values are not statistically significant they are part of an overall tendency for QUAD1 counties to adopt at a less rapid rate than QUAD4 counties, and for the ethnicity scores of QUAD1 counties to be higher than those of QUAD4 counties when forced adoption occurs.

Predictions cast for 1908 are the least accurate ones made. Accuracy is only 47.6 percent. Error rates are highest in QUAD4 where only 38.5 percent of all counties are correctly predicted. The 431 type II errors account for 81.9 percent of QUAD4 errors. In QUAD1 the proportion is almost identical, 82.7 percent of its errors are type II. Only 40.6 percent of QUAD1's counties are correctly predicted.

Overprediction is the rule on the other side of the ethnicity axis. 43.5 percent of QUAD3 counties are predicted dry, up 9.17 percent, but only 39.6 percent were observed dry, up 3.54 percent. This small divergence, 3.9 percent, helps keep the accuracy rate high in QUAD3 (where of the four quadrants it is highest) at 60 percent. In QUAD2 there are 286 errors, 249, 47.6 percent, of which are type I. Since 1892's predictions, the incidence of dry predictions climbed 9.58 percent. Simultaneously, the observed incidence dropped 7.33 percent. The loss of prohibition territory was actually more severe than the 7.33 percent figure shows because several QUAD2 counties
adopted after 1904. Of the 523 QUAD2 counties, 302 are predicted dry, but only 80 actually were. QUAD2 accuracy is 47.2 percent.

Aside from the errors caused by over and under prediction, 1908 is the most error prone year for a second reason. Of the 2499 counties in the analysis 1172 were dry in 1908. This represents 45.8 percent of all counties. The potential for error is greatest here because the model and reality can have completely separate estimates. The chance of error from this source declines in 1909 when 55 percent of the counties are dry.

National and regional historic events influenced state and local actions. The effect of these events are observed in the repeals by QUAD2 and QUAD3 counties. The evacuation from QUAD2, more so than from QUAD3, intensifies the effect of relocation first identified in 1884's predictions. The two are separate, yet complementary, events; a fact not identified during the spatial analysis.

5.45 Predictions for 1909 and 1910

The patterns of growth in 1909 and 1910 are very similar (Table 5.10 and Table 5.11). In 1909, for the first time, over 50 percent of the counties were dry. Between 1908 and 1909 over nine percent of the 2499 counties adopted, and between 1909 and 1910 another four percent adopted. Growth was evident in all quadrants, but it was most rapid in QUAD1 where the dry incidence was up 18.77 percent. Adoption was spurred by the activities of the Anti-Saloon League and statewide adoption in Oklahoma. Alabama, North Carolina, and Tennessee adopted in 1909 under the auspices of ASL leaders. Many Southern counties in other states took up local option; as did counties in every other
Table 5.10: Accuracy of the Socio-Economic Dimensions in 1909, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obs</td>
</tr>
<tr>
<td>wet wet</td>
<td>146</td>
</tr>
<tr>
<td>dry dry</td>
<td>80</td>
</tr>
<tr>
<td>wet dry*</td>
<td>274</td>
</tr>
<tr>
<td>dry wet**</td>
<td>23</td>
</tr>
</tbody>
</table>

percent predicted dry 67.69
percent observed dry 19.69
number observed dry 103
percent correct 43.21
n=523

percent predicted dry 30.79
percent observed dry 74.49
number observed dry 254
percent correct 37.24
n=341

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obs</td>
</tr>
<tr>
<td>wet wet</td>
<td>259</td>
</tr>
<tr>
<td>dry dry</td>
<td>187</td>
</tr>
<tr>
<td>wet dry</td>
<td>221</td>
</tr>
<tr>
<td>dry wet</td>
<td>113</td>
</tr>
</tbody>
</table>

percent predicted dry 52.31
percent observed dry 38.46
number observed dry 300
percent correct 57.18
n=780

percent predicted dry 47.37
percent observed dry 86.67
number observed dry 741
percent correct 48.54
n=855

* Type I error
** Type II error
Table 5.11: Accuracy of the Socio-Economic Dimensions in 1910, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry*</td>
</tr>
<tr>
<td>dry</td>
<td>wet**</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>73.61</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>21.80</td>
</tr>
<tr>
<td>number observed dry</td>
<td>114</td>
</tr>
<tr>
<td>percent correct</td>
<td>41.30</td>
</tr>
<tr>
<td>n=523</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry</td>
</tr>
<tr>
<td>dry</td>
<td>wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>60.26</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>44.87</td>
</tr>
<tr>
<td>number observed dry</td>
<td>350</td>
</tr>
<tr>
<td>percent correct</td>
<td>56.40</td>
</tr>
<tr>
<td>n=780</td>
<td></td>
</tr>
</tbody>
</table>

* Type I error
** Type II error
region, including the Mid-Atlantic for the first time because of ASL activity.

Between 1908 and 1909, there were 107 adoptions by QUAD4 counties, 45 by QUAD1, 23 by QUAD2 counties, and 51 by QUAD3 counties. Between 1909 and 1910, there were 32 adoptions in QUAD4, 19 in QUAD1, 9 in QUAD2, and 50 in QUAD3 counties. Many Southern counties adopted during this period. 85 percent of the QUAD1 adopters were from the South, the other 15 percent from the Border region. In Tennessee, statewide adoption forced the adoption of only four counties; three of these adopters were from QUAD1, the fourth from QUAD4. The mean ethnicity score of the three QUAD1 adopters was 1.03, and the score of the sole QUAD4 adopter was 0.240.

Forced adoption in Alabama involved 20 counties, seven from QUAD4 with a mean ethnicity score of 0.727, standard deviation 0.429, and twelve from QUAD1 with a mean ethnicity score of 2.241, standard deviation 0.214, and one QUAD3 county. In Alabama there were 31 QUAD1 counties, 31 QUAD4 counties, and one QUAD3 county. The effect of forced adoption was not as obvious as one might think, because the higher the percent black the higher the urbanness score. Southern counties that were over half black tended to reside in QUAD1. In Alabama, for example, QUAD4 counties were never over fifty percent black, with a mean of 20.8 percent based on the 1910 census. The Census of 1910 also reveals that only nine Alabama QUAD1 counties were less than half black. The mean percent black for Alabama QUAD1 counties was 58.4 percent. Fewer QUAD1 counties adopted on their own than QUAD4 counties, demonstrating the effect of the urbanness dimension both in terms of level of urbanization and percent black.
Two conflicting trends have emerged. During the mid 1890s the value of the variable NEGR890 was very high. Now we see that counties with large black populations are among the last to adopt. In the 1890s adoption was centered in Virginia and Mississippi and involved only 19 counties. Nine of the eleven counties from Mississippi were from QUAD1 as were three others, including one of the six from Virginia. This indicates that the high proportion black counties that adopted then were also QUAD1 counties. One influencing factor was the states where adoption was occurring during each of these periods. Political events inside Mississippi not present in Alabama, perhaps associated with the attack of Southern Conservatives on the voting rights of blacks, may have come to a head in the 1890s. Alternatively, activity by religious leaders may have influenced the black voters into voting in prohibition during the 1890s. Determination of which, if either of these hypotheses is correct is beyond the present research, but analysis of the social model has identified a point for future research that analysis of the spatial model did not.

As the two black and Protestant quadrants approach saturation, their predictive accuracy rates improve. In QUAD4 in 1909, 86.7 percent of the counties were dry and in 1910 this percentage climbs to 90.4 percent. Accuracy increases from 48.5 percent to 56.8 percent over the same two year period. In 1909 88.2 percent, 388, and in 1910 86.7 percent, 320, of all QUAD4 errors are type II. In QUAD1, which was 37.6 percent accurate in 1909 and 42.6 percent in 1910, the incidence of dry counties increased from 74.7 percent to 80.2 percent between 1909 and 1910. 85 percent, 181, and 84.2 percent, 165, of all QUAD1 errors were type I in 1909 and 1910, respectively. Accuracy in
these two quadrants is unavoidably on the increase. The declining ratio of type I to type II errors represents a real decline in the number of type I errors. The severe rate of underprediction coupled with the very high level of adoption means that nearly every county predicted dry in 1909 and 1910 was truly dry.

In QUAD3 accuracy rates are still on the decline, but the rate of decline is slowing. In 1909 QUAD3 accuracy drops to 57.2 percent from the 1908 level of 60 percent. In 1910 there is an additional loss of accuracy, 56.4 percent. The gap between the percent predicted and observed dry increases annually. In 1908 there are 339 counties predicted dry, but only 245 observed dry. In 1909 there are 408 QUAD3 counties predicted dry and 300 observed, and in 1910 it is 470 predicted and 350 observed dry. As the gap widens the proportion of errors shifts towards a preponderance of type I errors, 65.1 percent in 1908, 66.2 percent in 1909, and 67.6 percent in 1910 were type I.

In QUAD2 predictive accuracy continues on its downward trend. In 1909 it drops 4.4 percent from 1908's 47.6 percent and another 1.9 percent is lost in 1910. The percent of QUAD2 counties predicted dry increases from 47.6 percent in 1908, to 67.7 percent in 1909, to 73.6 percent in 1910. The observed percent dry, however, was much lower; 15.3 percent in 1908, 19.7 percent in 1909. In 1910, the QUAD2 observed incidence of dry counties, 21.8 percent, was only slightly less than the observed incidence for 1887. As the predicted dry number of counties continues to increase there is a noticeable decline in the number of type II errors, down four between 1908 and 1909 and another five between 1909 and 1910. Type I errors, on the other hand, increase dramatically in number because nearly all the predicted
conversion of wet to dry are premature. Type I errors increase dramatically as a result of the slow rate of growth and the high rate of predicted change. In 1910 over 55 percent, 289, of QUAD2's 523 counties are represented as type I errors. Nearly all predicted conversions in QUAD2 become type I errors.

The statistics for all four categories indicate that the model is failing, for the most part, because of the relocation event. The only group that is even closely approximated is QUAD3, and, even there the level of divergence is high and success rates barely above those that could be achieved from blind guessing.

5.46 Predictions for 1916, 1917 and 1918

The gap in predictions that occurs between 1910 and 1916 is caused by a period of repeal activity during the early teens. The peak occurred in 1911 when 64.1 percent of all counties were adopters. The trough occurred in 1914 when 58.7 percent were adopters. While this withdrawal is unnoticeable on any of the tables provided, territorial losses were confined, primarily, to QUAD1 and QUAD2 counties.

Why was there a decline in prohibition territory in the early teens? Only Alabama repealed prohibition during this period, and only eight of its counties became wet; seven from QUAD1 and one from QUAD3. The remainder of Alabama's counties retained local option. Otherwise, the source of withdrawal was the failure of counties to reaffirm their dry status at local option elections. Many states required the periodic reaffirmation, and in others local option elections could be called whenever an adequate number of names were attached to a petition. The repeal events are tentatively associated with the external
events in Europe. World War I, it is hypothesized, caused public turmoil similar to that caused by the Spanish-American War and two depressions. The turmoil resulted in a looser moral stance by enough people to negatively affect the prohibition movement. Alternatively, the decline in prohibition territory might be the result of a simple backlash effect, or the lagged consolidation of opposition forces among QUAD2 counties where expansion had begun one again. Histories of the movement dealing with this period focus on actors, not on the broader aspects of the movement. Detailed examination of events must be made before any of these hypotheses be accepted.

In 1916, the incidence of dry counties was 67.2 percent. In 1917, 75.2 percent of all counties were dry, and by 1918 there were 1995, 79.8 percent, adopters. Numerically the increase in dry counties was 197 between 1916 and 1917, and 107 between 1917 and 1918. Adoption status by quadrant reveals that 47.1 percent of all adopters in 1916 were located in QUAD4, where the highest proportion of dry counties existed, 94.0 percent (Table 5.12). Quad1 was 84.2 percent dry in 1916, QUAD3 60 percent dry, and QUAD2 was only 27.5 percent dry. By 1917, QUAD4 represented a smaller proportion of all dry counties, 42.8 percent. QUAD3, with 123 new adopters in 1917, gained the most ground on QUAD4, which only had 14 new adopters. Nearly eleven percent of the QUAD2 counties, 55, adopted in 1917. Only three QUAD1 counties adopted in 1917.

In 1917, or before, 95.1 percent of QUAD4's counties were adopters (Table 5.13). Adopter saturation is seen, therefore as the most important reason for the declining growth rate in QUAD4. The growth rate in QUAD1 was also slowing. Counties in QUAD1 were more highly
Table 5.12: Accuracy of the Socio-Economic Dimensions in 1916, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet wet</td>
<td>81</td>
</tr>
<tr>
<td>dry dry</td>
<td>116</td>
</tr>
<tr>
<td>wet dry*</td>
<td>298</td>
</tr>
<tr>
<td>dry wet**</td>
<td>28</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>79.16</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>27.53</td>
</tr>
<tr>
<td>number observed dry</td>
<td>144</td>
</tr>
<tr>
<td>percent correct</td>
<td>37.67</td>
</tr>
<tr>
<td>n=523</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet wet</td>
<td>110</td>
</tr>
<tr>
<td>dry dry</td>
<td>327</td>
</tr>
<tr>
<td>wet dry</td>
<td>197</td>
</tr>
<tr>
<td>dry wet</td>
<td>146</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>67.18</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>60.06</td>
</tr>
<tr>
<td>number observed dry</td>
<td>473</td>
</tr>
<tr>
<td>percent correct</td>
<td>56.03</td>
</tr>
<tr>
<td>n=780</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
</tr>
<tr>
<td>wet wet</td>
</tr>
<tr>
<td>dry dry</td>
</tr>
<tr>
<td>wet dry</td>
</tr>
<tr>
<td>dry wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
</tr>
<tr>
<td>percent observed dry</td>
</tr>
<tr>
<td>number observed dry</td>
</tr>
<tr>
<td>percent correct</td>
</tr>
<tr>
<td>n=855</td>
</tr>
</tbody>
</table>

* Type I error
** Type II error

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 5.13: Accuracy of the Socio-Economic Dimensions in 1917, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs pred freq percent</td>
<td>obs pred freq percent</td>
<td>obs pred freq percent</td>
<td>obs pred freq percent</td>
</tr>
<tr>
<td>wet wet</td>
<td>48</td>
<td>9.18</td>
<td>wet wet</td>
</tr>
<tr>
<td>dry dry</td>
<td>169</td>
<td>32.31</td>
<td>dry dry</td>
</tr>
<tr>
<td>wet dry*</td>
<td>276</td>
<td>52.77</td>
<td>wet dry</td>
</tr>
<tr>
<td>dry wet**</td>
<td>30</td>
<td>5.74</td>
<td>dry wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>85.09</td>
<td>percent predicted dry</td>
<td>58.36</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>38.05</td>
<td>percent observed dry</td>
<td>85.04</td>
</tr>
<tr>
<td>number observed dry</td>
<td>199</td>
<td>number observed dry</td>
<td>290</td>
</tr>
<tr>
<td>percent correct</td>
<td>41.49</td>
<td>percent correct</td>
<td>53.37</td>
</tr>
<tr>
<td>n=523</td>
<td></td>
<td>n=341</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs pred freq percent</td>
<td>obs pred freq percent</td>
</tr>
<tr>
<td>wet wet</td>
<td>43</td>
</tr>
<tr>
<td>dry dry</td>
<td>431</td>
</tr>
<tr>
<td>wet dry</td>
<td>141</td>
</tr>
<tr>
<td>dry wet</td>
<td>165</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>73.33</td>
</tr>
<tr>
<td>percent observed dry</td>
<td>78.40</td>
</tr>
<tr>
<td>number observed dry</td>
<td>596</td>
</tr>
<tr>
<td>percent correct</td>
<td>60.77</td>
</tr>
<tr>
<td>n=780</td>
<td></td>
</tr>
</tbody>
</table>

* Type I error
** Type II error

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
dispersed than counties in QUAD4, the minimum urbanness value was -1.01 and the maximum was 21.21. The very highly urban counties like New York, Queens, Philadelphia, and Cook were QUAD1 counties. These counties are widely dispersed and divergent from all other counties so the potential for contacting one of them is small, just as contact was unlikely with Western counties in the spatial model. In QUAD1 of the social model the wave front approaches the outlying highly urban counties, but it is prevented from simulating their adoption because these counties have very low contact probabilities. Contact probabilities are low because of the higher level of dispersion in their direction as compared to other directions. It is all right that highly urban counties are not predicted dry, because they do not, as a group, adopt until they are forced to by the adoption of larger, containing political units. Because some counties cannot be predicted dry QUAD1 experiences the same saturation effect at 85 percent dry that QUAD4 does at 95 percent dry.

The simulation closes with predictions for 1918. In 1918, there were 1995 adopters representing 79.8 percent of all counties used in this model (Table 5.14). 306 dry counties are in QUAD1, 247 in QUAD2, 624 in QUAD3, and 834 in QUAD4. There is a substantial difference in the ratio of adopters to total number of counties in each of the four quadrants. In QUAD1 89.7 percent were adopters, in QUAD2 47.2 percent, in QUAD3 81.7 percent, and in QUAD4 97.5 percent. These percentages indicate that the relocation event that occurred early in the study period and the repeal by several states with high QUAD2 and QUAD3 representation in the 1890s had continuous impact through national adoption. Further, they show how extensively rural and black
Table 5.14: Accuracy of the Socio-Economic Dimensions in 1918, by Quadrant

<table>
<thead>
<tr>
<th>Quadrant 2</th>
<th>Quadrant 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry*</td>
</tr>
<tr>
<td>dry</td>
<td>wet**</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>88.34</td>
</tr>
<tr>
<td>percent correct</td>
<td>45.51</td>
</tr>
<tr>
<td>n=523</td>
<td>percent correct</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quadrant 3</th>
<th>Quadrant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>obs</td>
<td>pred</td>
</tr>
<tr>
<td>wet</td>
<td>wet</td>
</tr>
<tr>
<td>dry</td>
<td>dry</td>
</tr>
<tr>
<td>wet</td>
<td>dry</td>
</tr>
<tr>
<td>dry</td>
<td>wet</td>
</tr>
<tr>
<td>percent predicted dry</td>
<td>79.10</td>
</tr>
<tr>
<td>percent correct</td>
<td>66.03</td>
</tr>
<tr>
<td>n=780</td>
<td>percent correct</td>
</tr>
</tbody>
</table>

* Type I error
** Type II error
Predictive accuracy continually improves during these three years. Accuracy for 1916 stands at 53.9 percent, up 2.4 percent from 1910's predictions. The gain in accuracy is, naturally, not equal for all quadrants. QUAD1 and QUAD4 gain in accuracy, 1.6 and 9.6 percent, respectively. QUAD2 and QUAD3 are 1.9 and 0.3 percent less accurate, respectively. The error ratios for the four quadrants reveal that underprediction is still the rule on the positive side of FACTOR2 and overprediction is confined to the negative side. The gains in accuracy in QUAD1 and QUAD4 coupled with type I to total error ratios of less than 0.2 provide further evidence for saturation in those quadrants.

QUAD2 was far from the point of saturation in 1916. Predictions for QUAD2 are, however, approaching the saturation point and predicted growth can be expected to slow before too much longer. The gap between the predicted and observed incidence of dry counties which peaked in 1910 at a 52 percent difference is just beginning to close in 1916, 51.7 percent. The accuracy of the model in QUAD2, however, is still declining. In 1916 QUAD2 accuracy reaches its trough at 37.7 percent accurate.

Accuracy is also lowest for QUAD3 in 1916. 60 percent of QUAD3's counties were adopters, and 67.2 percent are predicted dry. The accuracy level of 56 percent is still well above minimal accuracy, which for QUAD3 in 1916, given the observed and predicted incidence of dry territory, is 27.2 percent. Because both figures are above 50 percent and increasing, the potential for error is already on the decline.
For 1917 the model is 60.8 percent accurate, and accuracy is on the rise in every quadrant. Accuracy for QUAD1 is 53.3 percent, for QUAD2 it is 41.5 percent, 60.8 percent for QUAD3, and 76.1 percent for QUAD4. 79.7 percent of the counties in QUAD4 are predicted dry and only 36, 4.2 percent, are type I errors. Type I errors comprise only 15.8 percent of all QUAD4 errors. These figures are contrasted by the statistics from QUAD2 where 90.2 percent of all errors are type I. 85.1 percent of QUAD2's counties are predicted dry, yet only 38 percent were dry. 52.8 percent of all counties in QUAD2 are erroneously predicted dry and a mere 5.7 percent are type II errors.

Previous trends in prediction and reality continue in QUAD3 and QUAD1. Adoption in QUAD3 occurred rapidly, up 18.3 percent from 1916, there were 596 adopters in 1917. Accuracy, however, only climbs 4.7 percent, and there is a large shift in the type I to total ratio. The slow rise in accuracy results from there being only 37 fewer errors despite massive shifting among the various combinations. There are over 100 more counties predicted and observed dry, but there were 67 fewer wet and predicted wet counties. In QUAD1 the predicted incidence dry, still well below the observed incidence dry, increases considerably. In 1916 the quadrant is predicted to be 48.7 percent dry, but one year later it is predicted 58.4 percent dry. The gain occurs because only three QUAD1 counties adopted. The accuracy rate in QUAD1 is disturbingly close to the minimal accuracy value of 43.4 percent. Most of the blame for this low level of accuracy is assigned to the problems associated with continued underprediction.

Closing out this discussion of model errors we come to 1918 and the errors in its prediction. Due to overall saturation effects,
accuracy has been climbing during the last few predictions. Overall accuracy for 1918 is 67.5 percent, a value that is higher than that derived from the predictions of 1892. The error rate in QUAD1 declines substantially because the percent predicted dry gained 4.1 percent of the observed percent. Predictions for QUAD1 are 64.5 percent accurate. Most of the gain in accuracy comes from the conversion of type II errors into correct dry predictions. In QUAD2 the rate of predicted adoption began to show signs of saturation in 1918's predictions. Accuracy for QUAD2 is still below 50 percent, as is the percent observed dry. QUAD3 remains very consistent. 79.1 percent of its counties are predicted dry, and 81.7 percent were dry. The accuracy level, 66 percent, indicates that there are many errors despite this conformity. The error proportion is also very evenly split. 48.7 percent of all QUAD3 errors are type I. Of QUAD4's 855 counties 834 are dry in 1918. QUAD4's predictions are 84.7 percent accurate. This is excruciatingly close to minimal accuracy which is 84 percent. Underprediction in QUAD4 accounts for 112 of 133 errors.

Throughout the error analysis one multifaceted continuing trend is present. QUAD2 predictions well exceed the number of observed adoptions, while predictions for QUAD1 and QUAD4 fall well short of the mark. Unlike the preceding analysis where only this trend was identified, here it is possible to link the failure to several historic incidents, each of which increased the likelihood of adoption by QUAD1 and QUAD4 counties. First, religious leaders in QUAD4 and QUAD1 counties acted as change agents. Their actions increased the probability of adoption by counties in those quadrants. Second, there was massive repeal by QUAD2 and QUAD3 counties during the mid 1890s.
The repeals were sparked by several historic events, the most crucial of which were the Supreme Court's decision to uphold the Original Package law and massive European immigration to QUAD2 and QUAD3 counties. Third, a second period of repeals during the early teens saw the repeal of prohibition by many more urban than average counties. The model was not designed to accommodate shifts in the adopter profile. Because it only had base line information to operate from the high rate and observed organization of errors occurred.

5.5 Summary

Diffusion modeling in non-spatial dimensions provides the researcher with useful insights into the nature of adopters. Modeling in the socio-economic dimensions has shown that prohibition was more popular than expected amongst counties with higher than average concentrations of Negroes and less popular than predicted in urban counties with higher than average proportions of foreign born and Roman Catholics. The major reasons for this situation are the continuation of historic status, the frequent adoption of prohibition by other movements as a plank in their platforms, and the use of prohibition as a device to control segments of the population.

The socio-economic dimensions are created from 17 variables gathered from the Censuses of 1920, 1910, and 1900, as well as the Censuses of Religious Bodies for the years 1926, 1916, 1906, and 1896. If a county had been formed by 1896, it as was included in this aspect of the analysis. The variables used describe the major constituent elements of the population of each county including ancestry and religion of residents, the number of residents, the density of resi-
dents, the degree to which a county is urban, and the economic base of the county. Factor Analysis, a statistical technique which detects covariation between many variables simultaneously, was used to identify the two sources of covariance that explain the most variance. These major sources of covariance were then designated as variables, and each observation's value for the two variables was calculated. The calculated values for each observation become the observation's location within the matrix of the social dimensions. The variable generated by the factor analysis that explains the most variance is named 'urbanity' because the variables that are the major contributors to it are the indicators that define to what degree a county is urban. The higher the value of urbanity, the more urban the county. The other generated variable is called 'ethnicity' because the variables that contribute most to it define resident ancestry and religion. When the variable ethnicity is positive, the county has a higher proportion of county residents who are Negro than the mean county, and when the value of ethnicity is negative a higher proportion of county residents are foreign born and Roman Catholic than the mean county.

The socio-economic dimensions are accurate predictors of prohibition status 62.2 percent of the time. The most accurately predicted for year is 1884 at 92.1 percent accurate, and the model is least accurate using the social dimensions in 1908 when it is only 47.6 percent accurate. By dividing the distribution of factor scores into four groups, cartesian quadrants, it is discovered that different types of errors appear in each of the quadrants. Type II errors (dry counties that are predicted wet) are most common in QUAD1 and QUAD4 because the model consistently underpredicts the number of dry coun-
ties in each. Type I errors (wet counties that are predicted dry), on the other hand, are most common in QUAD2 where the model overpredicts the number of dry counties. County status is most accurately predicted in QUAD3 where, unlike the other three quadrants, accuracy never falls below 55 percent, and the number of type I and type II errors are very similar.

The distribution of model errors, when examined one prediction at a time, reveals that the socio-economic characteristics of adopting counties changed shortly after the Civil War. Counties adopting after the war tended to have a greater than average proportion of black residents. During the 1890s the the mean characteristics of counties that adopted prior to then changed because six Northern states repealed. The distribution also demonstrates that prohibition was a rural phenomena and that its ethnic focus changed over time depending who saw the movement as useful for furthering other goals.
CHAPTER SIX
THREE-DIMENSIONAL SIMULATION DIFFUSION MODELS

Analysis of the movement in the spatial dimensions demonstrates
the existence of a very strong tendency toward spatial, contagious
diffusion; the spatial model accounts for more of the variance in the
observed diffusion of prohibition than does the social model. Reform-
ing the model, using the social dimensions produces a less accurate
prediction than does the same model using the spatial dimensions.
Adopter similarity, in this instance, provides less basis for pre-
dicting diffusion than does adopter propinquity. The social dimen-
sions, through the use of error analysis, however, provide more in-
sight into growth patterns of the movement than do the spatial dimen-
sions by telling us what adopters were like. This chapter presents
one method for increasing still more the explanatory power of the
diffusion model.

The three-dimensional models presented here demonstrate a tech-
nique that blends the high level of accuracy noted in the spatial
dimensions with enhanced residual definition made possible by the
social dimensions. The operational explanation for the three-dimen-
sional model's choice of a county as an adopter is a function of both
the county's proximity and similarity to counties already observed and
predicted as dry. The three dimensional model assumes that counties
that are significantly different in the third dimension, however that
dimension is defined, but in close proximity are less likely to inter-
act than counties separated by the same distance that are more
similar. The variables chosen as the third dimensions for analysis in
this chapter achieve the desired blending by yielding accuracy levels intermediate to those of the two-dimensional models while better defining counties in the error analysis.

Two three-dimensional diffusion models, one based on adopter position in the urban hierarchy and one based ethnic-racial composition are compared in this chapter. The results of this comparison indicate that position in the urban hierarchy is slightly more explanatory than ethnic-racial composition when predicting the growth trends of prohibition. The innovation diffused in reverse order to that expected by the urban hierarchical model, despite the greater accuracy of the model in these three dimensions. The results also demonstrate that third dimensions other than position in the urban hierarchy are capable of accurately predicting growth patterns. Classic hierarchical diffusion, long considered the only example of diffusion in three dimensions, is reclassified as one of a multitude of modifiers of simple spatial expansion diffusion.

Presented first in this chapter are the program changes required to implement the three-dimensional model. There follow two examples of three-dimensional models that use the spatial dimensions and one of the socio-economic dimensions. The discovered qualities of three-dimensional models derived from these two runs are presented last.

6.1 Modifications Required for Three-Dimensional Modeling

Two modifications of program1's algorithm enable it to search for the nearest neighbors of each point in three dimensions when constructing the linked list data structure. The section of program1 that calculates distance between points requires modification to permit the
calculation of distance between the object point and the subject point in three dimensions. To accomplish this the solution of the Pythagorean theorem for the x-y dimensions is combined with the difference in the z dimension and the Pythagorean equation recalculated.

For program1 to decide where to place the next dry county a mechanism is provided that does so by means of finding the difference between each pair of the dimensions (x-x, y-y, z-z) and noting the sign of the differences. Sector identification, modified to reflect the integration of the third dimension, reduces the number of sectors from twelve to eight. Sector assignments are made based on the signs of the three subtractions used to calculate point-to-point distances. Each of the eight sectors for any given subject point is a cube, each of the eight cubes (forming a 2 x 2 x 2 sector block) share the subject point as a common corner. In all cases, the sign of zero difference between subject and object points is assumed positive. If, for example, the signs of all three point-to-point subtractions are positive, then the object point is assigned to sector one of the subject point. When all three signs are negative, the object point is assigned to sector eight.

Three-dimensional runs require only one modification to the simulation routine; the reduction from twelve to eight searches required to locate a random number because only eight sectors exist.

Alterations to the analysis program reflect the greater complexity of three-dimensional models. Indicators of overall accuracy are generated in the previous manner; the detailed analysis is, however, presented regionally for all years by dimensions. The observations are grouped into eight regions, the same regions as used in the two-
dimensional analyses. The third dimension is also categorized. Model accuracy and error distributions are discussed as functions of the regions and categorical values of the dimensions.

6.2 A Three-Dimensional Model Using Level of County Urbanization as the Third Dimension

The dimensions latitude, longitude, and urbanness (level of county urbanization), when combined in the three-dimensional simulation diffusion model correctly predict county prohibition status in 66.778 percent of the predicted cases. The correlation between percent of counties observed and predicted dry is 0.9977. The model using these three variables is distinct from hierarchical diffusion only because the most urban, yet spatially distant counties, are not directly connected in the data structure. The similarities are, otherwise, very great.

As in the two-dimensional models, accuracy rates vary greatly over time. In 1880 this model is 94.57 percent accurate. For 1881 and 1886 accuracy remains over 80 percent, but for 1887 it is only 72.52 percent. For the predictions of 1904 and 1915 is just over 57.1 percent. For the intervening years, 1907 and 1908, the model is less than 54 percent accurate. In 1914 the model is 55.24 percent accurate. Increasing accuracy observed in 1914 is delivered a setback in 1916 where the model's accuracy drops six percent to 49.21 percent correct. The break in rising accuracy in 1916 is related to the wave of recidivism that occurred in the early teens. For 1918 accuracy is 65.08 percent. Increased accuracy in the closing years is a result of saturation effects.
Analysis of model predictions and errors by region and urbanness scores indicate that the different rates of observed adoption by region and category are modeled much better within some groups than others. The analysis of this model's ability to predict prohibition status is divided into two sections. This division is established solely for explanatory purposes. The two sections of the analysis are diffusion in the urbanness dimension and diffusion in the spatial dimensions. Analysis based on the distribution of FACTOR1 scores supports the contention of several authors that the diffusion of prohibition was a true grass roots event. The regional analysis of the model reveals both changes and consistencies in the model's predictive ability as compared to the two-dimensional spatial model. The analysis of the three-dimensional model's ability to predict diffusion in the spatial dimensions shows that the inclusion of the third dimension causes a loss of accuracy in every region, relative to the accuracy of the model based on the spatial dimensions. The regions with the highest and lowest accuracies, the Mountain West (84.29 percent) and the Mid-Atlantic (51.03 percent), are both regions where adoption began after 1908, but that have nothing in common in any of the three dimensions (Table 6.1).

The high accuracy among the Western counties arises from their having been predicted dry throughout the study period due to their great subjective distance from dry territory. Errors in prediction occur in the West because the model fails to convert any counties, despite observed onset of adoption in the West in 1908. All errors in the West are, as one would expect by now, type I.

Simulated diffusion occurs rapidly in the Mid-Atlantic states...
Table 6.1: Accuracy of Three-Dimensional Model Using Latitude, Longitude and Factor one by Region and Factor one Scores

<table>
<thead>
<tr>
<th>Level of Urbanization Category</th>
<th>Region</th>
<th>less than 0</th>
<th>0 to 1</th>
<th>1 to 2</th>
<th>over 2</th>
<th>Reg. total</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td></td>
<td>79.545</td>
<td>80.729</td>
<td>53.509</td>
<td>63.726</td>
<td>68.386</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11)</td>
<td>(16)</td>
<td>(19)</td>
<td>(17)</td>
<td>(63)</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td></td>
<td>56.667</td>
<td>42.160</td>
<td>47.549</td>
<td>65.476</td>
<td>51.030</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(26)</td>
<td>(67)</td>
<td>(34)</td>
<td>(21)</td>
<td>(148)</td>
</tr>
<tr>
<td>NW Territory</td>
<td></td>
<td>62.379</td>
<td>60.221</td>
<td>64.491</td>
<td>55.000</td>
<td>61.633</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(232)</td>
<td>(141)</td>
<td>(32)</td>
<td>(10)</td>
<td>(415)</td>
</tr>
<tr>
<td>Plains</td>
<td></td>
<td>58.795</td>
<td>55.508</td>
<td>45.833</td>
<td>86.111</td>
<td>58.682</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(360)</td>
<td>(59)</td>
<td>(4)</td>
<td>(3)</td>
<td>(426)</td>
</tr>
<tr>
<td>South</td>
<td></td>
<td>71.621</td>
<td>63.741</td>
<td>61.728</td>
<td>53.333</td>
<td>70.809</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(508)</td>
<td>(284)</td>
<td>(27)</td>
<td>(5)</td>
<td>(824)</td>
</tr>
<tr>
<td>N-S Border</td>
<td></td>
<td>78.748</td>
<td>71.007</td>
<td>62.500</td>
<td>47.222</td>
<td>71.904</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(337)</td>
<td>(48)</td>
<td>(6)</td>
<td>(3)</td>
<td>(394)</td>
</tr>
<tr>
<td>Mountain</td>
<td></td>
<td>83.767</td>
<td>85.507</td>
<td>91.667</td>
<td></td>
<td>84.290</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(96)</td>
<td>(23)</td>
<td>(3)</td>
<td></td>
<td>(122)</td>
</tr>
<tr>
<td>Pacific</td>
<td></td>
<td>66.535</td>
<td>84.259</td>
<td>64.833</td>
<td>95.833</td>
<td>72.818</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(65)</td>
<td>(36)</td>
<td>(4)</td>
<td>(2)</td>
<td>(107)</td>
</tr>
<tr>
<td>Nation</td>
<td></td>
<td>69.450</td>
<td>58.749</td>
<td>57.801</td>
<td>63.388</td>
<td>66.778</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1634)</td>
<td>(674)</td>
<td>(129)</td>
<td>(61)</td>
<td>(2499)</td>
</tr>
</tbody>
</table>

n in parentheses

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
because no Mid-Atlantic county is more than seven edges from a county that is dry in 1876. Despite the overall low accuracy in the Mid-Atlantic states, predictions for highly urban counties (plus two on the urbanness scale) are vastly more accurate. The Mid-Atlantic counties, New York, Queens, Brooklyn, Bronx, Nassau, Erie, Philadelphia, Allegheny, New Castle, Hudson, and Camden, only one of which was really dry prior to 1919, are prevented from adopting in this simulation prior to 1919. The model correctly fails to designate these counties 'dry' because they are poorly linked (several of their pointers contain the null county) in the connectivity matrix and are therefore incapable of receiving a message.

Despite the increase in correct predictions made for the highly urban counties, accuracy is low throughout the Northeast, not just in the Mid-Atlantic states. In both New England and the Northwest Territory, recidivism, as in the two-dimensional models, is the major cause of low accuracy (68.39 percent and 61.63 percent respectively). In the Plains states, the level of accuracy is a mere 58.68 percent. The model mimics the observed adoption of Kansas in 1881 more poorly than does the two-dimensional model. It predicts the adoption of North Dakota as poorly as the spatial model does. In the South (70.81 percent) and in the Border states (72.82 percent), however, the model matches growth very well, both in terms of rate of growth and in the choice of adopter. Predictions for the Pacific Coast states are the second highest (72.82 percent) generated by the model. The same simulated events occur on the Pacific Coast as in the Mountain West, except that the degree of isolation from dry territory is increased. In reality, however, observed adoption on the coast predates adoption.
in the mountain states. This fact accounts for the slightly lower accuracy found among the coastal states when compared to the mountain states.

For the analysis of the model focused on FACTOR1 scores the urbanness variable for all 2499 counties is broken into four categories. The categories were chosen to reflect the skewed distribution of the factor analysis generated variable, FACTOR1. The ranges in value for the categories are: less than zero (n=1635), zero to one (n=674), one to two (n=129), and urbanness scores over two (n=61). Only one observation falls below negative one, the remainder of the observations in the first category fall between negative one and zero.

The model produces more accurate results for counties located on the extremes of the urbanness scale than for those in the central categories. The model is most accurate when predicting in the less-than-zero class (69.45 percent). Among the counties in the 0-to-1 class, the model accurately predicts status for 58.75 percent of the counties, and in the 1-to-2 class, it only correctly predicts 57.80 percent. For counties with urbanness scores in excess of two accuracy of predictions reaches 63.39 percent.

The poor separation of the moderately urban counties from the rural category is partially a function of the model's failure to identify the differential appeal of the movement. The simulated environment fails to reflect the observed rate and direction of change in the appeal of the movement. This is expressed as prematurely simulated adoption for counties in these categories. Type I errors account for 59.89 percent of all errors encountered in the zero to one category and 76.99 percent of errors in the one to two category. Dis-
counting the South and West, where no type I errors occur, the ratio of type I errors to all errors becomes 70.58 percent in the zero to one category and 89.78 percent in the one to two category. An explanation that fits the model's frame is that these counties are not marked in the third dimension as divergent enough, from their neighbors to reflect the observed degree to which they reject prohibition. In actuality, however, much of the failure can be attributed to events outside the model's operating environment and knowledge. Those self same events that have plagued previous models. Unlike the spatial model it is possible to recognize that a shift occurred in adopter social characteristics and location after the Civil War, not just the site of active adoption. The tendency was for rural counties to adopt at a higher rate after 1880 than prior to that date. The model failed to recognize this shift resulting in type I errors throughout the simulated northeastern United States.

The relatively high accuracy of the model in the +2 category is the result of the low level of connectivity of these counties in the data structure. Essentially, prohibition appealed little to these counties, and the model identified this fact. High accuracy in the low order counties, however, is the result of spatial trends, the isolation of Western counties, for example. Over half the rural counties are located in the South and Border regions. The high level of accurate predictions in these two regions in the rural category suggests that the model variables are important, relative indicators of the movements growth in, at least, part of the nation.

The movement was well on its way to expanding back into the North and into the West by 1914. At that time almost 70 percent of the low
order counties were dry as compared to 11.5 percent of the high order counties. This trend continued through 1918 when 89 percent of the counties in the >0 category were dry, 69.6 percent of the counties in the -1-to-0 range, 49.6 percent in the one to two range, and only 34.4 percent of the counties in the most urban category are dry. The direction of differential appeal verifies the grass roots orientation of the movement as described in Anti-Saloon League policy.

Error analysis of this model shows that the more urban the county the later it adopted prohibition. The prohibition movement flowed around high order places engulfing the low and eroding away the resistance of the encircled high order places. The percentage of counties dry in the various urbanness categories confirms this assertion. In the 1870s, the distribution of dry counties was relatively uniform across the range of FACTOR1. With relocation in the South and Border states, came accelerated adoption among rural counties. The uniform distribution of adopters across urban categories was disrupted by the rapid rate of growth among rural counties. Under the hypothesis of hierarchical diffusion this could not occur. Prohibition is, therefore, not a hierarchically diffused innovation.

Would the movement have relocated into the South were there not a large number of rural counties present that were pre-disposed towards adoption? This question essentially forms the distinction between the analyses in the spatial and social dimensions. We know that both spatial and social relocation occurred but which has precedence for explanatory purposes? Traditional spatial models have led into a "spacious cul-de-sac" where geographers are able to say where an adoption event occurred but not why (Blaikie, 1978). The identifica-
dition of adopter characteristics is a integral part of the social model, and in using it a better grasp of who adopts is achieved. While the question remains unanswerable, it is realized now that only by merging these models can we hope to adequately define adopters to the point where questions of why can be answered. While simply throwing more variables at the problem does not provide a solution, it is realized that the complex nature of human decision making requires multivariate explanation. The complex nature of human decision making requires multivariate explanations.

The addition of county urbanness as a third dimension has expanded our comprehension of the growth trends of the prohibition movement. The analysis of this three-dimensional model has revealed several facts relating to the diffusion of prohibition. 1) After a state repeals prohibition, the most urban counties readopted last. 2) The more urban a county, the later it adopts. 3) Prohibition diffused up the urban hierarchy. 4) The model is able to segregate, to a limited degree, counties that are unlikely to adopt and postpone their adoption even when all surrounding counties have adopted. 5) Proximity to dry territory is still the overriding factor in determining when a county adopts in the simulated world. 6) Relocation and growth are multivariate phenomena that require explanations of why adoption occurred and not mere descriptions of where adoption occurred.

Further, it has been demonstrated that a model that is an approximation of hierarchical diffusion as defined by economic geographers can be constructed using a Rapoport type data structure with three variables. It will next be demonstrated that hierarchical diffusion is not a special case, that other 'hierarchies' exist, and can be used
in diffusion models. These other hierarchies take the form of variables that are included with the spatial dimensions in a three-dimensional model.

6.3 A Three-Dimensional Model Using Ethnic-Racial Composition as the Third Dimension

The three-dimensional model using latitude, longitude, and ethnic-racial composition is 66.59 percent accurate. The correlation between observed and predicted percent accurate is 0.9985. Model accuracy differs by region by 26.4 percent. Fifteen of this model's predictions were matched to real years. The slow rate of predicted growth, represented by fifteen predictions as compared to twelve in the previous three-dimensional model, is achieved by a strong tendency for counties to recontact counties that were already predicted dry.

This model predicts the prohibition status in 1881 better than in any other year (88.15 percent). 1881 is, of course, the first predicted year. The trough in accuracy occurs when predictions are made for 1907 (54.56 percent). Accuracy reaches 62.13 percent in 1911 before dropping to 58.73 percent in 1916. At saturation in 1918 the model is 68.91 percent accurate. In terms of trends in accuracy this model is no different than any of its predecessors. Regional accuracy is greatest in the Mountain West (84.70 percent) and lowest among the Mid-Atlantic states (58.30 percent) in a now familiar pattern associated with regional proximity to counties that were dry in 1876. In all regions, except in the South and in New England, the ethnicity model produced higher accuracies than those generated by the urbanness three-dimensional model (Table 6.2).
Table 6.2: Accuracy of Three-Dimensional Model Using Latitude, Longitude and Factor two by Region and Factor two Scores

<table>
<thead>
<tr>
<th>Region</th>
<th>less than -1</th>
<th>-1 to 0</th>
<th>0 to 1</th>
<th>1 to 2</th>
<th>over 2</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>New England</td>
<td>63.604 (37)</td>
<td>73.867 (25)</td>
<td>93.333 (1)</td>
<td>68.042 (63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>60.833 (32)</td>
<td>58.554 (100)</td>
<td>45.238 (14)</td>
<td>96.667 (2)</td>
<td>58.299 (148)</td>
<td></td>
</tr>
<tr>
<td>NW Territory</td>
<td>71.097 (82)</td>
<td>59.665 (260)</td>
<td>72.394 (17)</td>
<td>40.000 (1)</td>
<td>100.000 (415)</td>
<td></td>
</tr>
<tr>
<td>Plains</td>
<td>67.616 (73)</td>
<td>60.484 (338)</td>
<td>68.444 (15)</td>
<td>61.975 (426)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>72.352 (22)</td>
<td>65.794 (47)</td>
<td>71.284 (353)</td>
<td>66.019 (321)</td>
<td>66.782 (81)</td>
<td>68.508 (824)</td>
</tr>
<tr>
<td>N-S Border</td>
<td>55.238 (7)</td>
<td>55.163 (51)</td>
<td>67.379 (316)</td>
<td>60.741 (18)</td>
<td>40.000 (2)</td>
<td>65.139 (394)</td>
</tr>
<tr>
<td>Mountain</td>
<td>92.929 (33)</td>
<td>81.667 (88)</td>
<td>78.571 (1)</td>
<td>84.699 (122)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific</td>
<td>90.991 (37)</td>
<td>69.609 (70)</td>
<td>77.003 (103)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nation</td>
<td>72.686 (323)</td>
<td>62.943 (979)</td>
<td>69.295 (771)</td>
<td>65.663 (340)</td>
<td>67.240 (86)</td>
<td>66.590 (2499)</td>
</tr>
</tbody>
</table>

$n$ in parentheses
The ethnicity dimension, which is almost normally distributed, is divided into five categories for analytical purposes. The first category is composed of counties that score below negative one in the ethnicity dimension. The second category holds counties with ethnicity scores between negative one and zero. Taken together, these two categories represent counties with ethnic-racial compositions with higher than average concentrations of Roman Catholics and foreign born residents. These counties also show a marked absence of Negro residents. The remaining three categories, representing counties having ethnicity scores greater than zero contain counties with increasingly higher proportions of Negro, Protestant, and native born residents. Counties with high proportions of both Negroes and foreign borns are located in the plus two category.

Ethnic values are regionally highly segregated (Table 6.2). This is the reason for the strong tendency for dry counties to be recontacted. Only in the South and Border states are there found substantial proportions of counties with ethnicity scores greater than zero. These two regions account for 86.77 percent of all counties in the zero to one category, 98.26 percent of the counties in the one to two category, and 96.51 percent of those in the plus two category. (To avoid the redundancy of a discussion of regional and ethnic trends the discussion here centers on ethnicity, with reference to spatial trends that ran counter to ethnic trends.)

Counties with ethnicity scores less than -1 are the most accurately predicted. Before 1916, less than 25 percent of the counties in this category had adopted prohibition. By 1918 less than 45 percent were adopters making this group the least likely to adopt. There
is little regional difference in the proportion of adopters within this category, but within it there is a great deal of difference in model accuracy by region. The few Southern counties in the less than -1 category were very slow to adopt and this fact was, to a large extent, correctly portrayed by the model. Since the Western counties were slow to adopt, and they were isolated in the simulation, predictive accuracy was characteristically high. The movement also did not appeal to Northeastern counties in this category but because of their proximity to dry territory they are predicted as having adopted. The higher density of counties in the less than -1 category accelerated simulated adoption in the Northeast. Model errors within this category are primarily type I. The model consistently overpredicts the percentage of adopters within the less than -1 category. Estimates are closely approximated prior to 1904, but after that time the observed and predicted rates begin to divergence.

The the negative one to zero ethnicity score category is the least well predicted group of counties. In this category, the model is 62.9 percent accurate. The isolation of counties in the Plains and Western regions from the movement's core area results in underprediction in those regions. As time goes by the proportion of type II errors in the West increases, but their number remains fairly constant in the Plains after 1900. In New England, the Mid-Atlantic states, and to a lesser extent in the Northwest Territory states, the opposite is true; premature prediction is the rule. Premature prediction, type I error, occurs because of the proximity of these counties to counties that were dry in 1876. Another cause for error in the Northeast is redicivism which affects accuracy in this category and region more.
than any other.

The model is substantially more accurate in the three positive categories than in the -1 to 0 category, but equally less accurate than the less than -1 category. The 0 to 1 category is 69.29 percent accurate, with its highest regional accuracy in New England (93.33 percent) followed by the Mountain West (78.57 percent) and the Northwest Territory (72.39 percent). All three regions combined, however, account for only 19 counties, or about one percent of all counties in the group. Suffolk County, Massachusetts is the sole New England county in the 0 to 1 category. As such, it is substantially different from its neighbors in the ethnicity dimension and it is effectively isolated from the them in the connectivity matrix. When simulated prohibition swept through New England, Suffolk County was the only one not converted. Northwest Territory counties in this ethnic-racial category are concentrated along the region's southern border and are spatially and ethnically akin to the counties of the Border states. This is one of the few instances outside the South where high accuracy is the result of accurate prediction of when the counties of the region adopt instead of high accuracy resulting from isolation or persistence. Accuracy among the Border states in this category is lower than might be expected due to an early tendency to mispredict before 1900, coupled with severe under-prediction afterwards. The same trend occurs in the South, but not to the same degree, as reflected by the relatively higher accuracy there.

Accuracy in the 1 to 2 category (65.66 percent) and the greater than 2 category (67.24 percent) is about the same. The South is the only region to contribute significantly to these categories, and
Southern accuracy is nearly identical for them both, approximately 66.3 percent. Southern counties with ethnicity scores greater than one were very likely to adopt during the 1880s and 90s. Because, however, of the ethnicity values of dry counties in 1876 the model selected these counties for adoption later than they really adopted. These type II errors represent, between 1904 and 1909, over 35 percent of all counties scoring over 1 on FACTOR2. Throughout the study period type II errors always account for over twenty percent of each category's counties. The results in these two categories indicate that the member counties were mistakenly separated from the rest of the Southern counties. Aside from those errors that have eroded the ability of all models, there are several sources of systematic error unique to this model.

The growth of the movement is related to the ethnic-racial dimension, although the relationship varies with time. The changing relationship results in a great number of errors. In the South 65.8 percent of all errors are type IIs. 89.4 percent of all errors in the Northwest Territory are type I. The Northwest Territory region include recidivous errors, which are a further reflection of the relocation event. Had relocation not occurred errors in each region would have been substantially fewer.

Before 1876, dry counties are concentrated in the less than 0 ethnicity categories, but starting shortly thereafter, dry territory begins to shift to the greater than 0 categories from where it spreads to counties in the remaining ethnic categories. This is nearly synonymous with the spatial interpretation that dry territory is concentrated in the Northeast prior to 1876, but that it spreads from the
South after it had relocated there around 1880. One must regard change in attitude as the cause for the relocation and not relocation as the cause for attitudinal modification.

On the whole separation of counties across this third dimension failed to increase predictive accuracy over that of the spatial dimensions alone. The one exception is the New England region. The problems in accurately modeling the diffusion of prohibition in all other models are re-encountered: accounting for the sectorial changes (relocation diffusion) that occurred during the 1890s, redicivism, and the lack of westward penetration of predicted dry territory.

6.4 Attributes of Three-Dimensional Models

To this point, the discussion has focused on the accuracy and error analysis of the two three-dimensional models. The emphasis here, however, is placed on the commonalities observed in the two implementations of the three-dimensional model. The common features of the two runs provide the best estimates available of what the general attributes and elements of three-dimensional models are.

The results of the two error analyses indicate that the information derived from the two three-dimensional model runs is equal to the information acquired from the two two-dimensional model runs. Each three-dimensional run, however, provides more information regarding the diffusion of prohibition than either of the two-dimensional models. The overlaying of three variables provides a better definition of the counties and the flow of the process than do two variables. If a researcher wishes to work with four variables, the three-dimensional approach is preferable because he can identify potential
adopter characteristics and process events in greater detail.

In multiple regression and analysis of variance the inclusion of additional variables often results in more variance being explained than the sum of the variance explained by each of the variables when used as the independent variable in a linear regression. This feature of these techniques is called a compound effect. That the amount of information derived from the two two-dimensional models is equivalent to that derived from the two three-dimensional models is proof that compound effects are not present in simulation diffusion models.

Model accuracy for the two three-dimensional models is intermediate compared to the accuracies observed in the two two-dimensional models. When data points are organized in these two sets of three dimensions, both networks are constructed such that they are less consistent with the network of observed adoption than the network of the spatial model. The different model network structures are very similar, and the relatively small fluctuations in model accuracy among all four runs have little effect on the differential explanatory power of each. Because each three-dimensional model explains more about the growth trends of prohibition than the spatial model, measures of model accuracy are of secondary importance. It must also be pointed out that not all third dimensions will cause a reduction in model accuracy. There exist several third dimensions that will increase the model's accuracy; they merely need to be identified.

What are the criteria that should be used to select a third dimension that will yield meaningful results? First of all, obviously, the researcher should know that the variable has some vital relationship with the innovation. Both third dimensions used here fit
this criterion as historical evidence has revealed. The third dimension chosen should be directional, that is, the early adopters should have third dimensional values clustered at one end of the dimensions range, and late adopters should progressively approach the other extreme. Neither of the dimensions used here can be classed as directional. The urbanness dimension is the more directional of the two, but because of the variable's skewed distribution, the expected clustering of rural county adoption near the beginning of the study period does not occur. Ethnic-racial composition is very non-directional owing to the relocation diffusion that occurs. The third dimension should not be highly correlated with the other two dimensions. When the third dimension is highly correlated with the first two dimensions, the pattern is reinforced more than modified. High correlation results in minimal reorganization of the network of connecting edges and little change in the flow of adopt signals. These observations suggest that a third dimension should be considered and accepted into the model based on the same criteria used to accept a new independent variable into a regression model.

The third dimension was observed to act as a barrier to diffusion in both three-dimensional models. The barrier effect occurs when a county is different from its two-dimensional neighbors in the third dimension. When the three-dimensional linkage structure is created that county is linked not with its two-dimensional neighbors but with other counties. Some observed changes in accuracy, as observed in the two-dimensional models, are accounted for by this fact. In the urbanness three-dimensional model, the scores of highly urban counties resulted in their being segregated from neighboring, less urban,
counties. The result was that urban counties were prevented from adopting in this simulation as early as they did in the two-dimensional spatial model. The high scoring counties in the ethnicity dimension were also prevented from adopting the innovation, but because of shifting patterns in the appeal structure of prohibition high scoring counties were not predicted dry as quickly as was appropriate. This aspect of the model can be manipulated by a researcher to segregate out certain individuals in a systematic manner.

The third dimension can also be used to build explicit, systematic barriers. Barriers have traditionally been implemented in two-dimensional models as walls beyond which no adoption can occur. Value ranges for the third dimension can be prevented from adopting regardless of their location in the other two dimensions. Simple isolation of individuals or the creation of pockets of resistance can be created by using the third dimension in this mode. During the period 1876 to 1904, for example, few counties with ethnicity scores under zero adopted, even though dry counties within this range existed.

To determine the impact of permitting adoption among these counties in the face of this knowledge, a model identical to that discussed in section 6.3, except for the inclusion of a barrier, was run. The barrier was initially positioned at zero and after each scanner generation the barrier was repositioned -.1 units on the ethnicity scale. This model increased accuracy by nearly six percent over the same model without the barrier. Use of this systematic barrier increased the accuracy of the model to what it had been in the spatial dimensions without the third dimension. Accuracy in the Southern and Border states declined, but improved accuracy in the remainder of the
nation, more than offset the loss. The model using the barrier did not increase comprehension, because barriers are built based on what is comprehended in the process, but not reflected in the model's operation. The presence of the barrier points out the amount of the error (about six percent of the total number of estimates made) associated with permitting adoption among counties scoring less than zero on the ethnicity variable in the basic three-dimensional model.

Three-dimensional models display a great deal of potential in future research. This potential is recognized not only in diffusion, but the connectivity matrix also has a potential use in social network theory. Of the four major failings of both three-dimensional models (recidivism, innovation relocation, random error, and failure to identify westward expansion), only the failure to identify westward expansion can be corrected given the definition and the constraints imposed on the model. The only method potentially available to adapt the model to reflect this growth is via the inclusion of a different third dimension. At least one third dimension exists that will perform this function without the use of barriers, but barriers can be used to achieve this goal with many other third dimensions. Barriers provide endless possibilities for improving accuracy, verifying suspicions regarding the direction and rate of growth among potential adopters, and identifying the effect of systematic errors in prediction.

The distribution of errors in three-dimensional models has more significance than the distribution of errors in two dimensions because the error analysis defines the adopters and non-adopters, at any time, more completely. In the examples given above, we know not only where they are, but what they are like in the third dimension as well.
Model accuracy is not necessarily improved by the inclusion of a third dimension and it may be affected detrimentally. Measures of model accuracy are less important, once a minimal threshold value is reached, than the analysis of the distribution of model errors over time to achieve an understanding of the diffusion of the innovation. The inclusion of a new variable creates a graph that, while still related to the previous graph, is different, and the determination of whether the new graph is a better representation of the observed is determined in the diffusion model. Finally, three-dimensional models point the way to defining that n-dimensional model that explains all variance, and for that reason, their development is an important step forward.
CHAPTER SEVEN

CONCLUSIONS

The comparative study of diffusion models has revealed much regarding the functioning and nature of the diffusion models examined. The analysis of the models has also permitted the identification of several previously unidentified growth trends in the spread of prohibition territory. The discussion in this concluding chapter is separated into two parts. The first section recapitulates information gained relating to the adopter tendencies behind the movement. The second section discusses the observed similarities and differences between the competing two-dimensional models and their synthesis, the three-dimensional model.

The growth of prohibition is unlike that of most innovations that have been studied by others. Adopters rejected and re-adopted the innovation at very high rates, resulting in both national and regional periodic growth and decline. Despite these periodic episodes of rapid acceptance and rejection, the movement grew slowly. Further, there were three levels of adopters studied: the county, the state, and the nation. County adoption could usually only occur after the state gave its permission for counties to determine their own status. The action of counties, however, had ramifications in the state legislatures and the nation's capital. On average, just under 62 percent of the counties in a state were dry the year before the state adopted. Eighty percent of all counties were dry when national prohibition was enacted.

The growth of prohibition was also different from technological
innovations because prohibition was not popular with all segments of society. This ambiguity became most obvious when everyone knew of the innovation, but most were slow to embrace it. As more groups adopted the middle class ethic, the innovation spread. Objections were politically overcome eventually in all places, but they were overcome first in those counties where the residents felt that the movement was consonant with their beliefs.

Each of the diffusion models contributes a unique view of the movement's spread. The merging of the interpretations provided by each perspective renders a much clearer image of what transpired than does any single model. Several trends that are distinct in one model are opaque in the others. The apparent importance of a factor in predicting the diffusion of prohibition differs greatly between models. The discussion of prohibition's growth trends must necessarily be centered around the four variables that were used in the models: latitude, longitude, urbanness, and ethnicity.

That prohibition was a rural phenomenon is readily apparent from examining the output from the models that included the urbanness variable. Rural people voted many a county seat dry over the objections of voters living there. When states repealed prohibition, the most urban counties re-adopted last. Regardless of which section of the country or group of demographic characteristics one discusses, the first counties to adopt were the more rural members of the group.

Why prohibition appealed more to rural folk than to those living in urban environs lies outside the scope of this inquiry. It was in part, however, a reactionary response. Drinking was, in many ways, a city thing. Our perception of drinking behavior is that it is a group
activity primarily done in bars and most bars are, and were, in towns. Farmers had little time for it, and they were under great social pressure to avoid the dangerous fluid. Support for prohibition has been identified as an indicator of membership, or aspiration to membership, in the middle class. During the late 1800s the middle class were not predominantly urban dwellers as they are today. Rather they resided in farming communities and on farms. A careful study of identifying exactly why prohibition appealed to rural America may be difficult to perform but it could add greatly to our knowledge of the fabric of nineteenth century American social geography. This aspect of the movement's growth requires substantial future research, however, before any hypotheses can be given solid support.

The concept of prohibition was a politicization of the moral suasion movement. Prohibition, just a little more radical than moral suasion, was the next step beyond the failure of the moral suasion movement. If people could not be convinced to abstain, then sobriety would be ordered! The politicization process is the major cause for the shifting demographic and spatial distributions of adopting counties.

The growth of prohibition was not as directional over time in the racial-ethnic or spatial variables. The activities of different religious sects and lobbies accounts for many of the changes, as represented by these variables, in adopter characteristics during the course of prohibition's period of expansion. Most of the activity of these groups was systematic in that they worked in specific types of communities, regions, or otherwise targeted a specific group. The results of each group's activity, however, when combined with the
results from all other groups generate a great deal of variance in the ethnic-racial, spatial, and temporal patterns of adoption.

The first political coalition prohibitionists formed was with the nativists, who were represented by the Know-nothing party. Prohibition and nativism had many sympathizers in common in the 1840s and 1850s. Each group drew its strength from native born and middle class men, and both groups were fearful of the changes in American society that immigrants caused. Prohibitionists could be found throughout the nation, but nativists were only present in areas receiving immigrants. The coalition, therefore, only formed in regions where both groups were present. The activities of the coalition resulted in adoption by a large portion of the nation prior to the Civil War. In demographic terms, prohibition passed in states receiving large numbers of immigrants, but in spatial terms, prohibition passed in Northern states.

In the 1890s, prohibitionists allied themselves with the populists. Populists supported the ambitions of the farmers and laborers, and as such populism was a working class movement. When the populists allied themselves with prohibitionists, they in essence, stated that to be middle class one must abstain. Large numbers of blacks voted for prohibition, indicating, perhaps, their desire for assimilation into mainstream America. During the last decade of the nineteenth century, growth of support for prohibition was slow in the South, but it was nil elsewhere. In ethnic terms, the counties that adopted prohibition during the 1890s were composed of Protestants, a very high proportion of whom were black.

Prohibitionists did not always tie their fate to that of other political factions. In 1872, prohibitionists struck out on their own
and formed the Prohibition party, which placed candidates on national ballots for over fifty years. During that time, however, the party never elected anyone to a major office. While many people voted for prohibition, few voted for prohibitionists. The lack of success of the Prohibition party demonstrated that, to achieve national prohibition, the movement's role had to be that of a plank and not a platform.

Prohibitionists formed two lobby organizations; the Women's Christian Temperance Union and the Anti-Saloon League. The WCTU and the ASL worked much as lobbyists do today. They organized marches, lobbied legislators, compiled statistics, published brochures, supported candidates whose views corresponded with theirs, and sabotaged their enemies. While the WCTU began in 1874 and the ASL in 1895, neither was a driving force in the movement before 1900. The ASL saw its main goal as territorial expansion; each action the league took was designed to yield a return in new dry counties and states. By starting small and consolidating territorial gains, they were able to achieve their ultimate goal. The WCTU was more of a social organization. Local WCTU chapters were mostly concerned with providing 'alternative' entertainment for young people and with keeping their own communities in 'proper' order. The national WCTU while supporting the local activities also sought to influence political leaders, but they had much less success than the ASL because WCTU members were disenfranchised. These two organizations linked the future of prohibition to individuals regardless of their party by backing any candidate from a major party who supported their cause. In this way they insured that the movement would not suffer when a party lost favor and
eventually filled house seats with their supporters.

Spatial and ethnic-racial dimensional relocation of the growth centers for prohibition are highly correlated with the distributions of the political organizations allied with prohibition. The willingness of the populace to support prohibition as a plank of several organizations throughout the century indicates the functional versatility of prohibition as a political weapon. In each occurrence, prohibitionists allowed their cause to be used as the weapon in another battle (native versus immigrant, black versus white, farmers versus city slickers). When one battle ended, or subsided, prohibition became the tool for another group. While prohibition's role in each of these fights brought some territorial gains it also weakened the movement when the allied movement fell into disfavor as evidenced by the resulting loss of territory after the collapse of the coalition (Figs. 1.3 and 1.4). The periods of decline noted in the movement's history are related to these alliances and to economic fluctuations (Clark, 1933).

Each alliance brought counties with different spatial, ethnic, racial, and religious characteristics into the dry category. The nativist link brought Northern counties with growing immigrant population into the fold. The religious link during the 1870s and 1880s with Southern denominations brought in Southern counties with moderate to low black populations. Ties with populists and their adversaries saw the adoption by Southern counties with large black populations. This despite the belief by many whites that blacks would fail to adopt prohibition because blacks, like Indians, could not control their drinking (Richard Beacon 1885; Walton and Taylor 1969; Walton 1972;
Woodward 1974, 60-63). After each alliance dissolved there was usually a short period of rapid territorial loss as counties that adopted under the coalition abandoned prohibition (Figs. 1.3 and 1.4).

Study of the relationships between prohibition and the other movements can reveal much about both American political processes and the interrelations of mass movements. Questions that come to mind include: Why did prohibitionists permit their cause to be used as a weapon in other, seemingly unrelated, political battles? To what degree were prohibition leaders also leaders in other social movements? Are lobby activities more effective than direct political confrontation? Do maps showing the intensity of affiliation with other political movements correlate well with maps of prohibition territory? Would the diffusion model be more accurate if information regarding the distribution of these other movements were included? Why were the depoliticizing activities of the ASL able to succeed?

While the questions posed above regarding the growth patterns of the prohibition movement are thought provoking, neither they nor the prohibition movement itself are the central focus of the current research. This experiment was designed to identify which of several diffusion models provides the most accurate results, the clearest, most comprehensive image of the adoption sequence, and to identify the factors that truly influenced the adoption process. A secondary goal was the discovery of modeling methods that improve accuracy and that provide greater comprehension of why a potential adopters adopt.

The comparative method provided a technique for identifying which model was most accurate. It was found that, in this case, the propinquity model is a substantially more accurate predictor of prohibition
status than the similarity model. The propinquity model is almost ten percent more accurate than the similarity model. This does not mean that for another innovation the same measures of similarity would be less predictive or that differently formulated measures of similarity would be less predictive of this innovation. Improved methods for identifying meaningful variables must be sought. In this case, variables that were hypothesized as good indicators in the social dimensions were found wanting. Improved accuracy without consideration for how and why adoption occurred is meaningless; an understanding of why the innovation was adopted is crucial.

Simply by running multiple diffusion models, a clearer image of the events surrounding the adoption process resulted than could have been obtained from any individual model. In this way, the experiment successfully demonstrated that diffusion researchers using spatial models should examine innovations from a wider perspective. For each model, spatial and social, we are able to provide both locational and demographic information regarding growth patterns, but only in terms of the defining set of dimensions. When spatial dimensions are used, questions regarding social patterns remain unformulated because aspects of the pattern in these dimensions are opaque to the spatial model, and vice versa. When the models are used together, however, the image, and thus the interpretation, of the events sharpen greatly. Patterns one might not have expected from simulating in one set of dimensions emerge in the other set.

Comparing the predictive ability of each set of dimensions does not determine the best set of dimensions to use. The spatial pattern is an after-the-fact display of where adoption occurred. Examinations
of why adoption occurred can only be made by stepping through the temporal pattern and asking why observation x adopted at the time it did. Only by asking this question of each observation does a trend become apparent to the researcher. In the study of technological innovations it is generally considered unnecessary to identify why people adopt. It is taken as given that new is desirable, and that upon receipt of information regarding the product adoption is automatic (once minor factors like economic and social resistance are overcome). Therefore the spatial model says that, those adjacent to an existing adopter will adopt next because they are in the best position to hear of the product.

In the case of social innovations the answer to why adoption occurs may be unrelated to the adopters position relative to others that have previously adopted. While receipt of information regarding the innovation precedes adoption, the adoption itself is more dependent of the characteristics of the observational unit. In models based in the social dimensions the various types of resistance are not a hindering factor that are unaccounted, but the driving force. Those with the least resistance adopt first, and those with the most last. Further, these characteristics change over time accounting for both recidivism and assimilation.

The prohibition movement provides an excellent example of the problems associated with using the simplistic notion that new is desirable as an explanation for adoption. Where the innovation is a reform that was heatedly debated, as prohibition was during the entire nineteenth century, use of the spatial model forces us to base explanations for model errors on supposed resistance. Relying on ad hoc
accounts of why adoption did not occur as expected in a given county is steeped in exceptionalism. Special cases are created to explain failure of historic fact to conform to predictions made under the spatial model. These exceptions become expressed in the models as errors and barriers. Significantly for one interested in actual historic landscapes, these barriers need have no standing beyond an ad hoc improvement of model accuracy. It is more appropriate to seek alternate variables for modeling, even if they provide less absolute predictive ability, than to create special cases for each failing.

Social models suggest two relationship between adopters and potential adopters. A conservative causal model states that an assumed relationship exists between similar counties identical in every way to spatially proximal counties. Because of the assumed relationship all the arguments invoked in spatial models hold. Adoption simply proceeds from counties very like the originator to those succeeding more dissimilar. This explanation yields an inadequate appraisal of human nature and of the facts regarding the innovation. Such a causal interpretation relies on the notion of resistance, but in this case the resistant dimensions are spatial, as represented by spatial coordinates of potential adopters. A more radical and more realistic, model suggests that counties adopt in an order dictated by assimilative modifications. The notion of pre-adaptation fits in this explanatory frame. Counties whose 'cultures' are most consistent with the movement adopt first, and as time passes, more and more counties modify their 'cultures' such that they can accept the innovation. As time passes counties less and less like the original adopters adopt. The modifications that permit adoption in each county remain unknown.
at this point, but each observation would adapt the innovation at its own rate and from its own starting point. In the case of prohibition, the factors that change bring the county's population into line with the culture of middle class America.

This mode of explanation does not rely on information flow; knowledge of the benefits and drawbacks of the innovation are presumed. The existence of such a cultural presumption underlies and justifies the mode of explanation. Investigating causality from this view with a slightly reformulated model, perhaps based in iterative discriminate analysis, may be a fruitful vein of future research.

The most significant indicator of the validity of the social model and its causal explanation lies in its providing more information about adopters than the spatial model. The social model's greater elegance lies in its definition of model residuals. The spatial model predicts the location of the next adopter from the location of a preceding adopter. When errors occur, therefore, we know only their spatial attributes because those variables alone are needed to drive the model. Using the social model we become aware of substantive qualities of the residuals because of the dimensional values used in the model. While explanations for residuals are by definition outside the explanatory frame of the models, explanations for errors of both models reside in social and historical factors such factors lie in very different data realm from the spatial variables, but conform to the social variables. The need to find explanation in a separate data source highlights a major shortcoming of spatial diffusion modeling.

A second and related problem with spatial models lies in the non-
systematic placement of barriers inserted in the model. Spatial barriers are placed in spatial dimensions as a result of the researcher's observing an anomalous spatial growth pattern. The researcher need not know the cause of the anomaly to make a correction. What appears as an anomaly in the spatial dimensions is, however, growth consistent with another, unrelated, but significant variable. A barrier that is placed systematically takes this variable and its values into account and relies on them for positioning the barrier. Instead of creating several spatial barriers to handle the same phenomenon in several locales, a single barrier based on that phenomenon can handle all occurrences. One simply needs to identify the variable in question and weight the probabilities of adoption according to its values.

The barrier used in the third model presented in chapter six demonstrates a systematic implementation of a barrier in addition to probability weighting for the third variable. In that case, it was slowly retracted allowing adoption to occur first among observational units that were more like the current adopters. The positioning of a barrier, when based on the values of another variable, has explanatory and causal meaning, but it remains an ad hoc attempt to increase accuracy without theoretical justification when applied without consideration of the causes of the 'deviant' spatial growth pattern. The construction of ad hoc barriers in response to an unidentified source of 'resistance' is a completely unscientific act. The need and justification for study in a wider perspective than just the spatial dimensions is again observed.

The expansion of a two-dimensional model to a three-dimensional may alleviate or reduce the need for barriers altogether. The three-
dimensional model incorporates a third dimension, rather than create special considerations for it. As such it is the optimal strategy for improving accuracy and explanation because no special cases are formed and no exceptions are granted. The difficulty resides in the identification of variables for this role. What were believed good selections for third dimensions in the models presented in chapter six were not. It proved easier, although less scientific, to invoke a barrier to increase accuracy. The invocation, however, did not improve understanding as the identification of an accuracy improving third dimension would have.

This discourse has led to a general improvement of diffusion models by expanding the scope of their operation and by demonstrating the advantages of using an increased number of variables. Prior to this work, the applications of the basic simulation diffusion model were limited to the spatial dimensions. The expansion of the model to encompass non-spatial dimensions has increased its functionality and versatility. To reflect this expansion of the model's realm of operation the term two-dimensional diffusion model replaces the term spatial diffusion model.

Three dimensional models are a natural outgrowth of two-dimensional models. Once I realized that any two variables could be substituted for the spatial variables in the standard diffusion model, it was a logical step to ask why three variables could not be combined in a model. Prior to this realization, other geographers had, of course considered the existence of a three dimensional model. Partially because of their orientation and their limiting themselves to the spatial dimensions, however, they perceived the three dimensional
model they developed, urban hierarchical diffusion, as a special case. In this special case, the "innovation potential of a center is a product of its position in the urban hierarchy and the force exerted on it by centers that have already adopted the innovation" (Berry 1971, 3).

Hierarchical diffusion has been seen as a 'short-circuiting' mechanism that overrides the general spatial trend. The concept of three-dimensional models is quite different. In three dimensional models, the observational units exist as points (or cells) in three dimensions. Probability of contact is defined in all three dimensions, not as two dimensions with a modification circuit traced over them as in the urban hierarchical model. The complete integration of the third dimension creates a more comprehensive and versatile model.

The re-identification of three dimensional models as a general type used in predicting the status of an observational unit instead of a special case is an important step forward. It was previously assumed that when hierarchical diffusion was in effect the innovation was transmitted from more urban places to less urban places. Examination of the social innovation, prohibition, reveals that the assumed direction of spread does not always hold. Future research will, I believe, demonstrate that social innovations spread from urban settings to rural ones as often as in the reverse direction. Study of prohibition also revealed that alternative hierarchies can predict the path of diffusion as well as the urban hierarchy. The two simple three-dimensional models run in chapter six were practically identical in their predictive ability, yet quite different paths were predicted in each case. While both three-dimensional models without barriers
were less accurate than the two-dimensional spatial model, there exist systems of three variables that are both more accurate in their predictions and provide more realistic explanations for why adoption events occur.

This research effort has revealed many facts regarding the growth of the prohibition movement, and fostered even more questions. In terms of the diffusion models, it demonstrates that the diffusion model as developed by Hagerstrand and Rapoport need not operate in the spatial dimensions alone. Applying the model to other dimensions allows researchers to observe growth patterns that would have gone unnoticed were the innovation modeled in the spatial dimensions alone. Further, the development of general three-dimensional models permits greater potential explanatory power, and incorporates a variant spatial diffusion model that was believed a special case. The development of three dimensional models also points to the future identification of n-dimensional models that account for even higher proportions of the observed variance. Much can yet be done to improve the model, and much needs to be done to determine whether the model deserves to be improved, redefined, or abandoned. Finally, this research has demonstrated once again the power of the comparative method for evaluating the performance of competing models.
BIBLIOGRAPHY


Anon. 1904. *Vermont Legislative Directory for 1904*.


Association Against the Prohibition Amendment. 1932. *Annual Report of the President*. Washington, D.C.


Clark, George Faber. 1888. History of the Temperance Reform in Massachusetts, 1813-1883. Boston: Clarke and Carruth.


Goodwin, Thomas A. 1883. Seventy-six Years' Tussle with the Traffic. Indianapolis: Carlon and Hollenbeck.


Hagerstrand, Torsten. 1965. *Aspects of the spatial structure of social communication and the diffusion of information*. *Papers, Regional Science Association* 16:


  National Home Rule Association.


  Lansing: Michigan Dry Campaign Committee.

  Commissioner of Excise of the State of New York*. New York:
  Wynkoop, Hallenbeck, Crawford.


Marble, Duane F. and Sophia Bowiby. 1968. Computer programs for the
  analysis of Hagerstrand type spatial diffusion models. *Northwestern
  University, Department of Geography, Technical Report* no. 9.


Marshall, Thomas K. 1915. *The First Six Months of Prohibition in
  Arizona*. Tucson, Arizona.

Martin, Asa. 1925. The temperance movement in Pennsylvania prior to
  the Civil War. *Pennsylvania Magazine of History and Biography*

McCarthy, Jeanne Bozze1l. 1980. *The Struggle for Sobriety:
  Protestants and Prohibition in Texas, 1919-1935*. El Paso:
  Texas Western Press.

McMichael, Claude E. 1951. Geographical reconnaissance of the area in
  Bienville and Webster Parishes, Louisiana. Master's thesis,
  Louisiana State University, Baton Rouge.


Mennel, S. J. 1969. Prohibition a sociological view. *Journal of
  American Studies* 3:159-75.

Monahan, M., ed. 1909. *A Textbook of True Temperance*. New York:
  United States Brewers Assoc.

  Charlottesville: University Press of Virginia.

Morrill, Richard L. 1968. Waves of spatial diffusion. *Journal of

Morrill, Richard L. 1970. The shape of diffusion in space and time.
  *Economic Geography* 46:259-68.


Scarp, Henry Anson. 1888. *King Alcohol in the Realm of King Cotton.* Atlanta: Blakely Printing.


Siglin, Daniel. 3 March 1983. Wyoming State Archives, Museums and Historical Department Personal Communication.


Stottard, Cora Frances. no date. *Fifteen Years of the Drink Question in Massachusetts*. Westerville, Ohio: American Issue Publishing.


Source listing Program1
(two-dimensional version with twelve 30 degree sectors)

DISTA:  PROCEDURE OPTIONS(MAIN);
#ifdef THIS PROGRAM CALCULATES THE DISTANCES BETWEEN POINTS AND
#endif
#ifdef IDENTIFIES THE 30 DEGREE SECTOR IN WHICH A POINT IS RELATIVE
#endif
#ifdef TO THE FIRST POINT. IT THEN FINDS THE CLOSEST POINT IN EACH
#endif
#ifdef SECTOR BY MEANS OF A SEQUENTIAL SEARCH. DISTANCES ARE THEN
#endif
#ifdef CONVERTED TO PROBABILITIES BY DIST**2 RULE.
#endif
DCL FAC012 FILE STREAM INPUT
DCL FILESTREAM ENV(FB BLKSIZE(6320) RECSIZE(80));
DCL 1 INFILE(2511),
3 ICPSR CHAR(7) VAR INIT('#'),
3 FACT1 FIXED(7,4) INIT(0),
3 FACT2 FIXED(7,4) INIT(0),
3 Y876 CHAR(1) VAR INIT(' ') IN,
1 OUTFILE,
3 ID FIXED(4) INIT(0),
3 SECTOR(12) FIXED(4) INIT(0),
3 PROB(12) FIXED(6,4) INIT(1),
LATM1 FLOAT INIT(0),
LATP1 FLOAT INIT(0),
LONM1 FLOAT INIT(0),
LONP1 FLOAT INIT(0),
I FIXED(4) INIT(0),
STOP FIXED INIT(2500),
K FIXED INIT(0),
L FIXED INIT(0),
M FIXED INIT(0),
TDIST FLOAT INIT(0),
J FIXED INIT(0);
DCL HDIST FLOAT INIT(0),
VDIST FLOAT INIT(0),
DIST2 FLOAT INIT(0),
ANGLE FLOAT INIT(0),
FAKER FIXED INIT(0),
ZONE FIXED INIT(0);
J=1;
DO J=1 TO STOP;
    GET FILE(FAC012) EDIT (ICPSR(J),FACT1(J),FACT2(J),Y876(J))
        (COL(6),A(7),COL(14),F(7,4),COL(22),F(7,4),COL(31),
A(1));
END;  /* J */
DO I=1 TO STOP;
    /* ESTABLISH SEARCH WINDOW +/-1 ABOUT POINT I */
    LATM1=FACT1(I)-5;
    LATP1=FACT1(I)+5;
LONM1 = FACT2(I) - 5;
LONP1 = FACT2(I) + 5;

/* CLEAR OUT SECTOR AND PROBABILITY FIELDS FOR THE NEXT OBS */
DO L=1 TO 12;
  SECTOR(L)=0;
  PROB(L)=25;
END;

DO K=1 TO STOP;
  /* CHECK IS POINT K WITHIN WINDOW */
  IF K=1 THEN K=K+1;
  IF FACT1(K) GT LATM1 THEN IF FACT1(K) LT LATP1 THEN IF FACT2(K) GT LONM1 THEN IF FACT2(K) LT LONP1 THEN CALL PLACE;
  END; /* K */
CALL PROBAB;
PUT SKIP EDIT(I,ICPSR(I),Y876(I),SECTOR(1),PROB(1),
  SECTOR(2),PROB(2),
  SECTOR(3),PROB(3),SECTOR(4),PROB(4),
  SECTOR(5),PROB(5),SECTOR(6),PROB(6),
  SECTOR(7),PROB(7),SECTOR(8),PROB(8),
  SECTOR(9),PROB(9),SECTOR(10),PROB(10),
  SECTOR(11),PROB(11),SECTOR(12),PROB(12))
  (X(1),F(4),X(1),A(7),X(1),F(4),F(6,4),
   F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),
   F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),
   F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),
   F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),
   F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),F(6,4),F(6,4))
END; /* I */
PLACE: PROCEDURE;
/* PROCEDURE IDENTIFIES SECTOR AND DISTANCE BETWEEN CENTRAL */
/* POINT I AND SURROUNDING POINTS. POINT BEING TESTED IS */
/* CLOSER THAN PREVIOUSLY ENCOUNTERED POINT IN THAT SECTOR*/
/* IT REPLACES THAT POINT IN THE FILE STRUCTURE OF OUTFILE*/
/* FOR POINT I */
   HDIST = FACT1(I)-FACT1(K);
   VDIST = FACT2(I)-FACT2(K);
   DIST2 = HDIST**2 + VDIST**2;
   IF DIST2 GT 25 THEN RETURN;
   IF DIST2 = 0 THEN DIST2 = .0001;
   ANGLE = ABS(VDIST) / (DIST2**.5);
   Faker = 1;
   IF ANGLE LE .5000 THEN Faker = 0;
   ELSE IF ANGLE GE .8660253 THEN Faker = 2;
SELECT;
   WHEN (HDIST GT 0 & VDIST LE 0) DO;
     IF Faker = 0 THEN Zone = 3;
     ELSE IF Faker = 1 THEN Zone = 2;
     ELSE IF Faker = 2 THEN Zone = 1;
   END;
   WHEN (HDIST LE 0 & VDIST LE 0) DO;
     IF Faker = 0 THEN Zone = 10;
     ELSE IF Faker = 1 THEN Zone = 11;
     ELSE IF Faker = 2 THEN Zone = 12;
END;

WHEN (HDIST GT 0 & VDIST GT 0) DO;
  IF FAKER=0 THEN ZONE=4;
  ELSE IF FAKER=1 THEN ZONE=5;
  ELSE IF FAKER=2 THEN ZONE=6;
END;

WHEN (HDIST LE 0 & VDIST GT 0) DO;
  IF FAKER=0 THEN ZONE=9;
  ELSE IF FAKER=1 THEN ZONE=8;
  ELSE IF FAKER=2 THEN ZONE=7;
END;

END; /* SELECT */

IF PROB(ZONE) GT DIST2
  THEN DO;
    PROB(ZONE)=DIST2;
    SECTOR(ZONE)=K;
  END;
END PLACE;

PROBAB: PROCEDURE;
/* PROCEDURE CALCULATES THE PROBABILITY OF CONTACT BASED ON 1 */
/* THE DISTANCE SQUARED FROM CENTRAL POINT FOR ALL 12 NEAREST */
/* POINTS. CONTENTS OF SUBSEQUENT SECTORS PROBABILITIES ARE */
/* STORED CUMULATIVELY IN THE PROBABILITY FIELDS */
  TDIST=0;
  DO M=1 TO 12;
    PROB(M)=25-PROB(M);
    TDIST=PROB(M)+TDIST;
    END;
  IF TDIST=0 THEN TDIST=.0001;
  DO M=1 TO 12;
    PROB(M)=PROB(M)/TDIST;
    END;
  DO M=2 TO 12;
    PROB(M)=PROB(M)+PROB(M-1);
    END;
  END PROBAB;
END DISTA;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
SIMULA: PROCEDURE OPTIONS(MAIN);
/* THIS PROCEDURE PERFORMS A SIMULATION RUN ANALOGOUS TO */
/* THE HAGERSTRAND SPATIAL DIFFUSION MODEL. METHOD OF CONTACT */
/* IS DERIVED FROM RAPOPORT. EACH OBS HAS 12 30 DEGREE FIELDS */
/* FOR CONTACT. PROBABILITY IS BASED ON DIST**2. M IS THE NUM */
/* OF YEARS SIMULATION IS FOR */
DCL 1 INFILE(0:2500),
  3 ID  FIXED(4) INIT(O),
  3 ICPSR CHAR(7) VAR INIT(''),
  3 Y876 CHAR(1) VAR INIT(''),
  3 SECTOR(12) FIXED(4) INIT((12)0),
  3 PROB(12) FIXED(6,4) INIT((12)0),
RESULT(2500) CHAR(45) VAR INIT(''),
STOP FIXED INIT(2499),
TEMP1(0:STOP) CHAR(1) VAR INIT(''),
NEX(0:STOP) CHAR(1) VAR INIT(''),
DRYIT FIXED INIT(O),
QQ FIXED INIT(O),
RR FLOAT INIT(501.77),
INT FLOAT INIT(O),
RANNUM FLOAT INIT(.55),
(I,J,K,L,M,N) FIXED INIT(O);
/*DCL DIRECT FILE STREAM INPUT; */
/* DO I=1 TO STOP;
 GET EDIT(ID(I),ICPSR(I),Y876(I),
  SECTOR(I,1),PROB(I,1),SECTOR(I,2),
  PROB(I,2),SECTOR(I,3),PROB(I,3),SECTOR(I,4),
  PROB(I,4),
  SECTOR(I,5),PROB(I,5),SECTOR(I,6),PROB(I,6),
  SECTOR(I,7),PROB(I,7),SECTOR(I,8),PROB(I,8),
  SECTOR(I,9),PROB(I,9),SECTOR(I,10),PROB(I,10),
  SECTOR(I,11),PROB(I,11),SECTOR(I,12),PROB(I,12))
  (COL(3),F(4),COL(8),
   A(7),A(1),X(1),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),
   F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),
   F(6,4),F(4),F(6,4));
TEMP1(I)=Y876(I);
NEX(I)=Y876(I);
RESULT(I)=Y876(I);
END;
DO M=1 TO 30;
QQ=0;
DO J=1 TO STOP;
  IF TEMP1(J) GE '1'
    THEN DO;
      CALL RANDOM;
      K=1;
      DO WHILE (K LT 13);
 Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
IF PROB(J,K) GT RANNUM
    THEN DO;
        DRYIT=SECTOR(J,K);
        NEX(DRYIT)=1;
        K=20;
        QQ=QQ+1;
        END;
    ELSE K=K+1;
    END; /* K */
END; /* J */
DO L=1 TO STOP;
    TEMP1(L)=NEX(L);
    RESULT(L)=RESULT(L)||NEX(L);
END; /* L */
PUT DATA(QQ);
END; /* M */
DO N=1 TO STOP;
    PUT SKIP EDIT(ICPSR(N),RESULT(N))(X(1),A(7),X(2),A(45));
END;

RANDOM: PROCEDURE;
    RR=RR*899;
    RR=RR/32767;
    INT=FLOOR(RR);
    RANNUM=RR-INT;
    IF RR LE .01 THEN RR=DRYIT;
    IF RANNUM=0 THEN RANNUM=.99;
    IF RANNUM GT .99 THEN RANNUM=.99;
    END RANDOM;
END SIMULA;
Source listing Program3
(version non-specific)

//DISSE1 JOB,'SECCHRST',TIME=(,30)
/*JOBPARM R=45,L=10
//STEP1 EXEC FLXCEG
//SYSIN DD *
MERGE: PROCEDURE OPTIONS(MAIN);
/* THIS PROGRAM MERGES THE RESULTS OF INDIVIDUAL SIMULATION */
/* RUNS TO DERIVE THE OMEGA PATTERN. 5 FILE ARE READ IN */
/* AND ADDED CELL BY CELL TO A SINGLE OUTPUT FILE. IF ALL 5 */
/* RUNS OF SIMUL SAID TIME-PLACE WAS DRY OUTPUT CELL HOLDS */
/* A 5. 4 SAID HOMES 4 ETC. */

DCL FR1(45) CHAR(1) VAR INIT(''),
FR2(45) CHAR(1) VAR INIT(''),
FR3(45) CHAR(1) VAR INIT(''),
FR4(45) CHAR(1) VAR INIT(''),
FR5(45) CHAR(1) VAR INIT(''),
ICPSR CHAR(7) VAR INIT(''),
OMEGA(45) FIXED(1) INIT(0),
J FIXED INIT(0),
(I,K,L)

DCL FRES1 FILE STREAM INPUT;
DCL FRES2 FILE STREAM INPUT;
DCL FRES3 FILE STREAM INPUT;
DCL FRES4 FILE STREAM INPUT;
DCL FRES5 FILE STREAM INPUT;

DO I=1 TO 2499;
GET FILE(FRES1) EDIT(ICPSR,(FR1(K) DO K=1 TO 45))
(COL(2),A(7),COL(11),45(A(1)));
GET FILE(FRES2) EDIT((FR2(K) DO K=1 TO 45))
(COL(11),45(A(1)));
GET FILE(FRES3) EDIT((FR3(K) DO K=1 TO 45))
(COL(11),45(A(1)));
GET FILE(FRES4) EDIT((FR4(K) DO K=1 TO 45))
(COL(11),45(A(1)));
GET FILE(FRES5) EDIT((FR5(K) DO K=1 TO 45))
(COL(11),45(A(1)));

DO J=1 TO 45;
OMEGA(J)=0;
IF FR1(J)='1' THEN OMEGA(J)=OMEGA(J)+1;
IF FR2(J)='1' THEN OMEGA(J)=OMEGA(J)+1;
IF FR3(J)='1' THEN OMEGA(J)=OMEGA(J)+1;
IF FR4(J)='1' THEN OMEGA(J)=OMEGA(J)+1;
IF FR5(J)='1' THEN OMEGA(J)=OMEGA(J)+1;
END;
PUT EDIT(ICPSR,(OMEGA(L) DO L=1 TO 45))
(COL(5),A(7),45(F(1)));
END;
END MERGE;

//GO.FRES1 DD DSN=Y13RPS1.FRES1,UNIT=DISK,_VOL=SER=ACA003,
// DCE=(LRECL=80,BLKSIZE=2960,RECFM=FB),DISP=SHR
//GO.FRES2 DD DSN=Y13RPS1.FRES2,UNIT=DISK,VOL=SER=ACA003,
//  DCB=(LRECL=80,BLKSIZE=2960,RECFM=FB),DISP=SHR
//GO.FRES3 DD DSN=Y13RPS1.FRES3,UNIT=DISK,VOL=SER=ACA003,
//  DCB=(LRECL=80,BLKSIZE=2960,RECFM=FB),DISP=SHR
//GO.FRES4 DD DSN=Y13RPS1.FRES4,UNIT=DISK,VOL=SER=ACA003,
//  DCB=(LRECL=80,BLKSIZE=2960,RECFM=FB),DISP=SHR
//GO.FRES5 DD DSN=Y13RPS1.FRES5,UNIT=DISK,VOL=SER=ACA003,
//  DCB=(LRECL=80,BLKSIZE=2960,RECFM=FB),DISP=SHR
Two-dimensional primary (spatial) analysis run

//SPACE JOB , 'SECHRIST', TIME=(, 30)
//*JOBPARM R=45, L=5
//*ROUTE PRINT RMT4
//STEP1 EXEC SAS
//INFO1 DD DSN=Y13RPS1.DISS.TOPC,UNIT=DISK,VOL=SER=ACA003,
   // DISP=SHR
//INFO2 DD DSN=Y13RPS1.DISS.SPATOMEQ,UNIT=DISK,VOL=SER=ACA003,
   // DISP=SHR
DATA PROH; INFILE INFO1;
    INPUT FIPS $ 2-6 YR FOUND 7-9
    LATDEG 10-11 LATMIN 13-14 LONDEG 15-17 LONMIN 19-20
    ICPSR $ 21-27 (Y876-Y919)(1.) XRDY 72-75 REGION 21;
    LATIT= LATDEG + LATMIN/60;
    LONGI= LONDEG + LONMIN/60;
    IF Y876 GT 1 THEN Y876=1;
    IF Y877 GT 1 THEN Y877=1;
    IF Y878 GT 1 THEN Y878=1;
    IF Y879 GT 1 THEN Y879=1;
    IF Y880 GT 1 THEN Y880=1;
    IF Y881 GT 1 THEN Y881=1;
    IF Y882 GT 1 THEN Y882=1;
    IF Y883 GT 1 THEN Y883=1;
    IF Y884 GT 1 THEN Y884=1;
    IF Y885 GT 1 THEN Y885=1;
    IF Y886 GT 1 THEN Y886=1;
    IF Y887 GT 1 THEN Y887=1;
    IF Y888 GT 1 THEN Y888=1;
    IF Y889 GT 1 THEN Y889=1;
    IF Y890 GT 1 THEN Y890=1;
    IF Y891 GT 1 THEN Y891=1;
    IF Y892 GT 1 THEN Y892=1;
    IF Y893 GT 1 THEN Y893=1;
    IF Y894 GT 1 THEN Y894=1;
    IF Y895 GT 1 THEN Y895=1;
    IF Y896 GT 1 THEN Y896=1;
    IF Y897 GT 1 THEN Y897=1;
    IF Y898 GT 1 THEN Y898=1;
    IF Y899 GT 1 THEN Y899=1;
    IF Y900 GT 1 THEN Y900=1;
    IF Y901 GT 1 THEN Y901=1;
    IF Y902 GT 1 THEN Y902=1;
    IF Y903 GT 1 THEN Y903=1;
    IF Y904 GT 1 THEN Y904=1;
    IF Y905 GT 1 THEN Y905=1;
    IF Y906 GT 1 THEN Y906=1;
    IF Y907 GT 1 THEN Y907=1;
    IF Y908 GT 1 THEN Y908=1;
    IF Y909 GT 1 THEN Y909=1;
    IF Y910 GT 1 THEN Y910=1;
DATA SPA; INFILE INFO2;
INPUT STATE 1-2 ICPSR $ 1-7 (W1-W20)(1.);
IF W1 GE 3 THEN X1=1; ELSE X1=0;
IF W2 GE 3 THEN X2=1; ELSE X2=0;
IF W3 GE 3 THEN X3=1; ELSE X3=0;
IF W4 GE 3 THEN X4=1; ELSE X4=0;
IF W5 GE 3 THEN X5=1; ELSE X5=0;
IF W6 GE 3 THEN X6=1; ELSE X6=0;
IF W7 GE 3 THEN X7=1; ELSE X7=0;
IF W8 GE 3 THEN X8=1; ELSE X8=0;
IF W9 GE 3 THEN X9=1; ELSE X9=0;
IF W10 GE 3 THEN X10=1; ELSE X10=0;
IF W11 GE 3 THEN X11=1; ELSE X11=0;
IF W12 GE 3 THEN X12=1; ELSE X12=0;
IF W13 GE 3 THEN X13=1; ELSE X13=0;
IF W14 GE 3 THEN X14=1; ELSE X14=0;
IF W15 GE 3 THEN X15=1; ELSE X15=0;
IF W16 GE 3 THEN X16=1; ELSE X16=0;
IF W17 GE 3 THEN X17=1; ELSE X17=0;
IF W18 GE 3 THEN X18=1; ELSE X18=0;
IF W19 GE 3 THEN X19=1; ELSE X19=0;
IF W20 GE 3 THEN X20=1; ELSE X20=0;
PROC SORT DATA=PROH; BY ICPSR; RUN;
PROC SORT DATA=SPA; BY ICPSR; RUN;
DATA ALL;
MERGE PROCH(IN=INC) SPA(IN=INE); BY ICPSR;
IF INC & INE;

IF Y878=. THEN Z878=.;
IF Y878=0 & X1=0 THEN Z878=0;
ELSE IF Y878=1 & X1=1 THEN Z878=1;
ELSE IF Y878=0 & X1=1 THEN Z878=2;
ELSE IF Y878=1 & X1=0 THEN Z878=3;

IF Y881=. THEN Z881=.
IF Y881=0 & X2=0 THEN Z881=0;
ELSE IF Y881=1 & X2=1 THEN Z881=1;
ELSE IF Y881=0 & X2=1 THEN Z881=2;
ELSE IF Y881=1 & X2=0 THEN Z881=3;

IF Y885=. THEN Z885=.
IF Y885=0 & X3=0 THEN Z885=0;
ELSE IF Y885=1 & X3=1 THEN Z885=1;
ELSE IF Y885=0 & X3=1 THEN Z885=2;
ELSE IF Y885=1 & X3=0 THEN Z885=3;
IF Y887=1 THEN Z887=.5
IF Y887=0 & X4=0 THEN Z887=0;
ELSE IF Y887=1 & X4=1 THEN Z887=1;
ELSE IF Y887=0 & X4=1 THEN Z887=2;
ELSE IF Y887=1 & X4=0 THEN Z887=3;

IF Y891=. THEN Z891=;.
IF Y891=0 & X5=0 THEN Z891=0;
ELSE IF Y891=1 & X5=1 THEN Z891=1;
ELSE IF Y891=0 & X5=1 THEN Z891=2;
ELSE IF Y891=1 & X5=0 THEN Z891=3;

IF Y904=. THEN Z904=;.
IF Y904=0 & X6=0 THEN Z904=0;
ELSE IF Y904=1 & X6=1 THEN Z904=1;
ELSE IF Y904=0 & X6=1 THEN Z904=2;
ELSE IF Y904=1 & X6=0 THEN Z904=3;

IF Y907=. THEN Z907=;.
IF Y907=0 & X7=0 THEN Z907=0;
ELSE IF Y907=1 & X7=1 THEN Z907=1;
ELSE IF Y907=0 & X7=1 THEN Z907=2;
ELSE IF Y907=1 & X7=0 THEN Z907=3;

IF Y908=. THEN Z908=;.
IF Y908=0 & X8=0 THEN Z908=0;
ELSE IF Y908=1 & X8=1 THEN Z908=1;
ELSE IF Y908=0 & X8=1 THEN Z908=2;
ELSE IF Y908=1 & X8=0 THEN Z908=3;

IF Y909=. THEN Z909=;.
IF Y909=0 & X9=0 THEN Z909=0;
ELSE IF Y909=1 & X9=1 THEN Z909=1;
ELSE IF Y909=0 & X9=1 THEN Z909=2;
ELSE IF Y909=1 & X9=0 THEN Z909=3;

IF Y910=. THEN Z910=;.
IF Y910=0 & X10=0 THEN Z910=0;
ELSE IF Y910=1 & X10=1 THEN Z910=1;
ELSE IF Y910=0 & X10=1 THEN Z910=2;
ELSE IF Y910=1 & X10=0 THEN Z910=3;

IF Y916=. THEN Z916=;.
IF Y916=0 & X11=0 THEN Z916=0;
ELSE IF Y916=1 & X11=1 THEN Z916=1;
ELSE IF Y916=0 & X11=1 THEN Z916=2;
ELSE IF Y916=1 & X11=0 THEN Z916=3;

IF Y917=. THEN Z917=;.
IF Y917=0 & X15=0 THEN Z917=0;
ELSE IF Y917=1 & X15=1 THEN Z917=1;
ELSE IF Y917=0 & X15=1 THEN Z917=2;
ELSE IF Y917=1 & X15=0 THEN Z917=3;
IF Y918=. THEN Z918=.;
IF Y918=0 & X18=0 THEN Z918=0;
ELSE IF Y918=1 & X18=1 THEN Z918=1;
ELSE IF Y918=0 & X18=1 THEN Z918=2;
ELSE IF Y918=1 & X18=0 THEN Z918=3;
PROC FREQ; TABLES W1 W2 W3 W4 W5 W6 W7 W8 W9 W10
  W11 W12 W13 W14
  W15 W16 W17 W18 W19 W20;
PROC FREQ;BY REGION;TABLES W1 W2 W3 W4 W5 W6 W7 W8
  W9 W10 W11 W12 W13 W14
  W15 W16 W17 W18 W19 W20;
PROC FREQ;BY REGION;
TABLES Z879 Z881 Z885 Z887 Z891 Z904 Z907 Z908
  Z909 Z916 Z917 Z918;
Source listing Program 1
(three-dimensional version with eight cubic sectors)

//DISSER JOB 'SECHRIST', TIME=(5), REGION=1024K
//JOBPARM R=45, I=30
//ROUTE PRINT RMT4
//STEP EXEC PLIXCEG
//SYSIN DD *
DISTA: PROCEDURE OPTIONS(MAIN);
/*****************************/
/* THIS PROCEDURE CALCULATES THE DISTANCES BETWEEN POINTS IN */
/* THREE DIMENSIONS. THE DISTANCES BETWEEN EACH PAIR OF TWO */
/* POINTS ARE SUBTRACTED, AND BASED ON THE SIGNS OF THE DIFF */
/* THE RELATIONSHIP IS ASSIGNED TO ONE OF EIGHT ZONES. */
/* THESE ZONES ARE SEGMENT OF A SPHERE 5 UNITS IN RADIUS */
/* ONCE THE CLOSEST POINTS IN EACH OF THE 8 ZONES HAS BEEN */
/* IDENTIFIED THE DISTANCES ARE CONVERTED TO PROBABILITIES */
/* OF CONTACT BY THE DIST**2 RULE. THE VALUE OF THE THIRD */
/* DIMENSION, THE EIGHT POINTERS AND PROBABILITIES ARE PRINT */
/*****************************/
DCL THREED FILE STREAM INPUT
ENV(FB BLSIZE(6320) RECSIZE(80));
DCL 1 INFILE(2511),
3 ICPSR CHAR(7) VAR INIT(''),
3 LATIT FIXED(7,4) INIT(0),
3 LONGI FIXED(7,4) INIT(0),
3 FACT1 FIXED(7,4) INIT(0),
3 Y876 CHAR(1) VAR INIT(''),
1 OUTFILE,
3 ID FIXED(4) INIT(0),
3 SECTOR(8) FIXED(4) INIT(0),
3 PROB(8) FIXED(6,4) INIT(1),
LATM1 FLOAT INIT(0),
LATP1 FLOAT INIT(0),
LONM1 FLOAT INIT(0),
LONP1 FLOAT INIT(0),
FACM1 FLOAT INIT(0),
FACP1 FLOAT INIT(0),
I FIXED(4) INIT(0),
STOP FIXED INIT(2500),
K FIXED INIT(0),
L FIXED INIT(0),
M FIXED INIT(0),
TDIST FLOAT INIT(0),
J FIXED INIT(0);
DCL HDIST FLOAT INIT(0),
VDIST FLOAT INIT(0),
WDIST FLOAT INIT(0),
DIST2 FLOAT INIT(0),
ZONE FIXED INIT(0);
J=1;
DO J=1 TO STOP;
GET FILE(THREED) EDIT (ICPSR(J),LATIT(J),LONGI(J), FACT1(J),Y876(J))
(COL(6),A(7),COL(14),F(7,4),COL(22), F(7,4),COL(30),F(7,4),COL(50),A(1));
END; /* J */
DO I=1 TO STOP;
/* ESTABLISH SEARCH WINDOW +/-5 UNITS ABOUT POINT I */
LATM1=LATIT(I)-5;
LATP1=LATIT(I)+5;
LONM1=LONGI(I)-5;
LONP1=LONGI(I)+5;
FACH1=FACT1(I)-5;
FACP1=FACT1(I)+5;
/* CLEAR OUT SECTOR AND PROBABILITY FIELDS FOR THE NEXT OBS */
DO L=1 TO 8;
SECTOR(L)=0;
PROB(L)=25;
END;
DO K=1 TO STOP;
/* CHECK IS POINT K WITHIN WINDOW */
IF K=I THEN K=K+1;
IF LATIT(K) > LATM1 THEN IF LATIT(K) < LATP1 THEN IF LONGI(K) > LONM1 THEN IF LONGI(K) < LONP1 THEN CALL PLACE;
END; /* K */
CALL PROBAB;
PUT SKIP EDIT(I,ICPSR(I),Y876(I),SECTOR(1),PROB(1), SECTOR(2),PROB(2), SECTOR(3),PROB(3),SECTOR(4),PROB(4), SECTOR(5),PROB(5),SECTOR(6),PROB(6), SECTOR(7),PROB(7),SECTOR(8),PROB(8))
(X(1),F(4),X(1),A(7),A(1),X(1),F(4),F(6,4), F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4), F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4))
END; /* I */
PLACE: PROCEDURE;
/* PROCEDURE IDENTIFIES SECTOR AND DISTANCE BETWEEN CENTRAL */
/* POINT I AND SURROUNDING POINTS. POINT BEING TESTED IS */
/* CLOSER THAN PREVIOUSLY ENCOUNTERED POINT IN THAT SECTOR */
/* IT REPLACES THAT POINT IN THE FILE STRUCTURE OF OUTFILE */
/* FOR POINT I */
HDIST=LATIT(I)-LATIT(K);
VDIST=LONGI(I)-LONGI(K);
WDIST=FACT1(I)-FACT1(K);
DIST2=HDIST**2+VDIST**2+WDIST**2;
IF DIST2 > 25 THEN RETURN;
IF DIST2=0 THEN DIST2=.0001;
SELECT;
WHEN (HDIST>0 & VDIST>0 & WDIST>0) ZONE=1;
WHEN (HDIST>0 & VDIST>0 & WDIST<=0) ZONE=2;
WHEN (HDIST>0 & VDIST<=0 & WDIST>0) ZONE=3;
WHEN (HDIST>0 & VDIST<=0 & WDIST<=0) ZONE=4;
WHEN (HDIST<=0 & VDIST>0 & WDIST>0) ZONE=5;
WHEN (HDIST<=0 & VDIST>0 & WDIST<=0) ZONE=6;
WHEN (HDIST<=0 & VDIST<=0 & WDIST>0) ZONE=7;
WHEN (HDIST<=0 & VDIST<=0 & WDIST<=0) ZONE=8;
END;  /* SELECT */
IF PROB(ZONE) > DIST2
THEN DO;
   PROB(ZONE)=DIST2;
   SECTOR(ZONE)=K;
END;
END PLACE;
PROBAB: PROCEDURE;
/* PROCEDURE CALCULATES THE PROBABILITY OF CONTACT BASED ON 1 */
/* THE DISTANCE SQUARED FROM CENTRAL POINT FOR ALL 12 NEAREST */
/* POINTS. CONTENTS OF SUBSEQUENT SECTORS PROBABILITIES ARE */
/* STORED CUMULATIVELY IN THE PROBABILITY FIELDS */
   TDIST=0;
   DO M=1 TO 8;
      PROB(M)=25-PROB(M);
      TDIST=PROB(M)+TDIST;
   END;
   IF TDIST=0 THEN TDIST=.0001;
   DO M=1 TO 8;
      PROB(M)=PROB(M)/TDIST;
   END;
   DO M=2 TO 8;
      PROB(M)=PROB(M)+PROB(M-1);
   END;
END PROBAB;
END DISTA;
//GO.THREED DD DSN=Y13RPS1.DISS.DIMEN4,VOL=SER=ACA001,UNIT=DISK,
//   DCB=(LRECL=80,BLKSIZE=6320,RECFM=FB),DISP=SHR
Source listing Program2
(three-dimensional version with eight cubic sectors)

//DISSEr JOB 'SECHRIST', TIME=(1), REGION=756K
//JObPARm R=45, LINES=10
//ROUTE PRINT RMT4
//STEP1 EXEC PLXCEG
//SYsIN DD *
SIMULA: PROCEDURE OPTIONS(MAIN);
/* THIS PROCEDURE PERFORMS A SIMULATION ANALOGOUS TO */
/* THE HAGERSTRAND SPATIAL DIFFUSION MODEL. METHOD OF CONTACT */
/* IS DERIVED FROM RAPOPORT. EACH OBS HAS 8 30 DEGREE FIELDS */
/* FOR CONTACT. PROBABILITY IS BASED ON DIST**2. M IS THE NUM */
/* OF YEARS SIMULATION IS FOR */
DCL 1 INFILE(0:2500),
   3 ID FIXED(4) INIT(0),
   3 ICPSR CHAR(7) VAR INIT(''),
   3 Y876 FIXED(4) INIT((8)0),
   3 SECTOR(8) FIXED(4) INIT((8)0),
   3 PROB(8) FIXED(6,4) INIT((8)0),
   RESULT(2500) CHAR(45) INIT(''),
   STOP FIXED INIT(2499),
   TEMP(0:STOP) CHAR(1) INIT(''),
   NEX(0:STOP) CHAR(1) VAR INIT(''),
   DRYIT FIXED INIT(0),
   QQ FIXED INIT(0),
   RR FLOAT INIT(756.3984),
   INT FLOAT INIT(0),
   RANNUM FLOAT INIT(.55),
   (I,J,K,L,M,N) FIXED INIT(0);
DCL DIRECT FILE STREAM INPUT;
DO L=1 TO STOP;
   GET FILE(DIRECT) EDIT(ID(I),ICPSR(I),Y876(I),
      SECTOR(I,1),PROB(I,1),SECTOR(I,2),
      PROB(I,2),SECTOR(I,3),PROB(I,3),SECTOR(I,4),
      PROB(I,4),
      SECTOR(I,5),PROB(I,5),SECTOR(I,6),PROB(I,6),
      SECTOR(I,7),PROB(I,7),SECTOR(I,8),PROB(I,8)
      (COL(2),F(4),COL(7),
      A(7),A(1),X(1),F(4),F(6,4),F(4),F(6,4),F(4),F(6,4),F(4),
      F(4),F(6,4),F(4),F(6,4),COL(2),F(4),F(6,4),F(4),
      F(6,4),F(4),F(6,4));
   TEMP(I)=Y876(I);
   NEX(I)=Y876(I);
   RESULT(I)=Y876(I);
   END;
/* M IS THE NUMBER OF YEARS TO SEQUENCE THROUGH */
/* J IS THE NUMBER OF OBSERVATIONS */
/* K IS THE NUMBER OF THE CURRENT POINTER */
DO M=1 TO 40;
   QQ=0;
   DO J=1 TO STOP;
IF TEMP1(J) >= '1'
    THEN DO;
        CALL RANDOM;
        K=1;
        DO WHILE (K < 8);
            IF PROB(J,K) > RANNUM
                THEN DO;
                    DRXIT=SECTOR(J,K);
                    NEX(DRXIT)='1';
                    K=20;
                    QQ=QQ+1;
                    END;
                ELSE K=K+1;
                END; /* K */
            END; /* J */
    END;

DO L=1 TO STOP;
    TEMP1(L)=NEX(L);
    RESULT(L)=RESULT(L) || NEX(L);
    END; /* L */

PUT DATA(QQ);
    END; /* M */

DO N=1 TO STOP;
    PUT SKIP EDIT(ICPSR(N),RESULT(N))(X(1),A(7),X(2),A(45));
    END;

RANDOM: PROCEDURE;
    RR=RR*899;
    RR=RR/32767;
    INT=FLOOR(RR);
    RANNUM=RR-INT;
    IF RR < .01 THEN RR=DRXIT;
    IF RANNUM=0 THEN RANNUM=.99;
    IF RANNUM > .99 THEN RANNUM=.99;
    END RANDOM;
END SIMULA;
//GO.DIRECT DD DSN=Y13RPS1.DIST3D.FACT1,UNIT=DISK, VOL=SER=ACA003,
//    DCB=(LRECL=80,BLKSIZEx=2960,RECFM=FB), DISP=SHR
Analysis
three-dimensional primary (latitude, longitude and ethnicity)
analysis run

//DISSER JOB 'SECHRIST', TIME=(30)
*/JOBPARM
*/ROUTE PRINT RMT4
/STEP1 EXEC SAS
/STEP1 DD DSN=Y13RPS1.DISS.TOPC,UNIT=DISK, VOL=SER=ACAO03,
// DCB=(LRECL=80, BLKSIZE=6400, RECFM=FB), DISP=SHR
/STEP2 DD DSN=Y13RPS1.OMEO3D.FACT2, UNIT=DISK, VOL=SER=ACAO03,
// DCB=(LRECL=80, BLKSIZE=6320, RECFM=FB), DISP=SHR
/STEP3 DD DSN=Y13RPS1.DISS.DIMEN4, UNIT=DISK, VOL=SER=ACAO03,
// DCB=(LRECL=80, BLKSIZE=6320, RECFM=FB), DISP=SHR
/STEP4 EXEC SAS
DATA DIMEN; INFILE STEP3;
INPUT ICPSR $ LATITUDE LONGITUDE URBAN ETHNIC;
IF ETHNIC < -1 THEN ETHX=' LT-2';
IF ETHNIC GT -1 AND ETHNIC LE 0 THEN ETHX=' -1 TO 0 ';
IF ETHNIC GT 0 AND ETHNIC LE 1 THEN ETHX=' 0 TO 1 ';
IF ETHNIC GT 1 AND ETHNIC LE 2 THEN ETHX=' 1 TO 2 ';
IF ETHNIC GT 2 THEN ETHX=' GT 2 ';
DATA PROB; INFILE STEP1;
INPUT ICPSR $ 21-27 (Y876-Y919)(1.);
IF Y876 GT 1 THEN Y876=1;
IF Y877 GT 1 THEN Y877=1;
IF Y878 GT 1 THEN Y878=1;
IF Y879 GT 1 THEN Y879=1;
IF Y880 GT 1 THEN Y880=1;
IF Y881 GT 1 THEN Y881=1;
IF Y882 GT 1 THEN Y882=1;
IF Y883 GT 1 THEN Y883=1;
IF Y884 GT 1 THEN Y884=1;
IF Y885 GT 1 THEN Y885=1;
IF Y886 GT 1 THEN Y886=1;
IF Y887 GT 1 THEN Y887=1;
IF Y888 GT 1 THEN Y888=1;
IF Y889 GT 1 THEN Y889=1;
IF Y890 GT 1 THEN Y890=1;
IF Y891 GT 1 THEN Y891=1;
IF Y892 GT 1 THEN Y892=1;
IF Y893 GT 1 THEN Y893=1;
IF Y894 GT 1 THEN Y894=1;
IF Y895 GT 1 THEN Y895=1;
IF Y896 GT 1 THEN Y896=1;
IF Y897 GT 1 THEN Y897=1;
IF Y898 GT 1 THEN Y898=1;
IF Y899 GT 1 THEN Y899=1;
IF Y900 GT 1 THEN Y900=1;
IF Y901 GT 1 THEN Y901=1;
IF Y902 GT 1 THEN Y902=1;
IF Y903 GT 1 THEN Y903=1;
IF Y904 GT 1 THEN Y904=1;
IF Y905 GT 1 THEN Y905=1;
IF Y906 GT 1 THEN Y906=1;
IF Y907 GT 1 THEN Y907=1;
IF Y908 GT 1 THEN Y908=1;
IF Y909 GT 1 THEN Y909=1;
IF Y910 GT 1 THEN Y910=1;
IF Y911 GT 1 THEN Y911=1;
IF Y912 GT 1 THEN Y912=1;
IF Y913 GT 1 THEN Y913=1;
IF Y914 GT 1 THEN Y914=1;
IF Y915 GT 1 THEN Y915=1;
IF Y916 GT 1 THEN Y916=1;
IF Y917 GT 1 THEN Y917=1;
IF Y918 GT 1 THEN Y918=1;
IF Y919 GT 1 THEN Y919=1;
DATA CEN; INFILE INFO2;
INPUT REGION 5 STATE 5-6 ICPSR $ 5-12 §13 (X1—X40)(1.);
IF STATE=34 THEN REGION=5;
IF X1 GE 3 THEN X1=1; ELSE IF X1 NE . THEN X1=0;
IF X2 GE 3 THEN X2=1; ELSE IF X2 NE . THEN X2=0;
IF X3 GE 3 THEN X3=1; ELSE IF X3 NE . THEN X3=0;
IF X4 GE 3 THEN X4=1; ELSE IF X4 NE . THEN X4=0;
IF X5 GE 3 THEN X5=1; ELSE IF X5 NE . THEN X5=0;
IF X6 GE 3 THEN X6=1; ELSE IF X6 NE . THEN X6=0;
IF X7 GE 3 THEN X7=1; ELSE IF X7 NE . THEN X7=0;
IF X8 GE 3 THEN X8=1; ELSE IF X8 NE . THEN X8=0;
IF X9 GE 3 THEN X9=1; ELSE IF X9 NE . THEN X9=0;
IF X10 GE 3 THEN X10=1; ELSE IF X10 NE . THEN X10=0;
IF X11 GE 3 THEN X11=1; ELSE IF X11 NE . THEN X11=0;
IF X12 GE 3 THEN X12=1; ELSE IF X12 NE . THEN X12=0;
IF X13 GE 3 THEN X13=1; ELSE IF X13 NE . THEN X13=0;
IF X14 GE 3 THEN X14=1; ELSE IF X14 NE . THEN X14=0;
IF X15 GE 3 THEN X15=1; ELSE IF X15 NE . THEN X15=0;
IF X16 GE 3 THEN X16=1; ELSE IF X16 NE . THEN X16=0;
IF X17 GE 3 THEN X17=1; ELSE IF X17 NE . THEN X17=0;
IF X18 GE 3 THEN X18=1; ELSE IF X18 NE . THEN X18=0;
IF X19 GE 3 THEN X19=1; ELSE IF X19 NE . THEN X19=0;
IF X20 GE 3 THEN X20=1; ELSE IF X20 NE . THEN X20=0;
IF X21 GE 3 THEN X21=1; ELSE IF X21 NE . THEN X21=0;
IF X22 GE 3 THEN X22=1; ELSE IF X22 NE . THEN X22=0;
IF X23 GE 3 THEN X23=1; ELSE IF X23 NE . THEN X23=0;
IF X24 GE 3 THEN X24=1; ELSE IF X24 NE . THEN X24=0;
IF X25 GE 3 THEN X25=1; ELSE IF X25 NE . THEN X25=0;
IF X26 GE 3 THEN X26=1; ELSE IF X26 NE . THEN X26=0;
IF X27 GE 3 THEN X27=1; ELSE IF X27 NE . THEN X27=0;
IF X28 GE 3 THEN X28=1; ELSE IF X28 NE . THEN X28=0;
IF X29 GE 3 THEN X29=1; ELSE IF X29 NE . THEN X29=0;
IF X30 GE 3 THEN X30=1; ELSE IF X30 NE . THEN X30=0;
IF X31 GE 3 THEN X31=1; ELSE IF X31 NE . THEN X31=0;
PROC SORT DATA=PROH; BY ICPSR;
PROC SORT DATA=CEN; BY ICPSR;
DATA ALL;
OPTIONS NOCENTER;
MERGE PROM(IN=INC) CEN(IN=INE); BY ICPSR;
IF INC & INE;

PROC SORT DATA=ALL; BY ICPSR;
PROC SORT DATA=DIMEN; BY ICPSR;
DATA NEWALL;
MERGE ALL(IN=EXC) DIMEN(IN=EXD); BY ICPSR;
IF EXC & EXD;
IF Y881=. THEN Z881=.;
IF Y881=0 & X2=0 THEN Z881=0;
   ELSE IF Y881=1 & X2=1 THEN Z881=1;
   ELSE IF Y881=0 & X2=1 THEN Z881=2;
   ELSE IF Y881=1 & X2=0 THEN Z881=3;
IF Y883=. THEN Z883=.;
IF Y883=0 & X3=0 THEN Z883=0;
   ELSE IF Y883=1 & X3=1 THEN Z883=1;
   ELSE IF Y883=0 & X3=1 THEN Z883=2;
   ELSE IF Y883=1 & X3=0 THEN Z883=3;
IF Y886=. THEN Z886=.;
IF Y886=0 & X4=0 THEN Z886=0;
   ELSE IF Y886=1 & X4=1 THEN Z886=1;
   ELSE IF Y886=0 & X4=1 THEN Z886=2;
   ELSE IF Y886=1 & X4=0 THEN Z886=3;
IF Y887=. THEN Z887=.;
IF Y887=0 & X5=0 THEN Z887=0;
   ELSE IF Y887=1 & X5=1 THEN Z887=1;
   ELSE IF Y887=0 & X5=1 THEN Z887=2;
   ELSE IF Y887=1 & X5=0 THEN Z887=3;
IF Y904=. THEN Z904=.;
IF Y904=0 & X7=0 THEN Z904=0;
   ELSE IF Y904=1 & X7=1 THEN Z904=1;
   ELSE IF Y904=0 & X7=1 THEN Z904=2;
   ELSE IF Y904=1 & X7=0 THEN Z904=3;
IF Y906=. THEN Z906=.;
IF Y906=0 & X8=0 THEN Z906=0;
   ELSE IF Y906=1 & X8=1 THEN Z906=1;
   ELSE IF Y906=0 & X8=1 THEN Z906=2;
   ELSE IF Y906=1 & X8=0 THEN Z906=3;
IF Y907=. THEN Z907=.;
IF Y907=0 & X9=0 THEN Z907=0;
   ELSE IF Y907=1 & X9=1 THEN Z907=1;
   ELSE IF Y907=0 & X9=1 THEN Z907=2.
ELSE IF Y907=1 & X9=0 THEN Z907=3;

IF Y908=. THEN Z908=.
IF Y908=0 & X10=0 THEN Z908=0;
ELSE IF Y908=1 & X10=1 THEN Z908=1;
ELSE IF Y908=0 & X10=1 THEN Z908=2;
ELSE IF Y908=1 & X10=0 THEN Z908=3;

IF Y909=. THEN Z909=.
IF Y909=0 & X12=0 THEN Z909=0;
ELSE IF Y909=1 & X12=1 THEN Z909=1;
ELSE IF Y909=0 & X12=1 THEN Z909=2;
ELSE IF Y909=1 & X12=0 THEN Z909=3;

IF Y910=. THEN Z910=.
IF Y910=0 & X14=0 THEN Z910=0;
ELSE IF Y910=1 & X14=1 THEN Z910=1;
ELSE IF Y910=0 & X14=1 THEN Z910=2;
ELSE IF Y910=1 & X14=0 THEN Z910=3;

IF Y911=. THEN Z911=.
IF Y911=0 & X15=0 THEN Z911=0;
ELSE IF Y911=1 & X15=1 THEN Z911=1;
ELSE IF Y911=0 & X15=1 THEN Z911=2;
ELSE IF Y911=1 & X15=0 THEN Z911=3;

IF Y916=. THEN Z916=.
IF Y916=0 & X16=0 THEN Z916=0;
ELSE IF Y916=1 & X16=1 THEN Z916=1;
ELSE IF Y916=0 & X16=1 THEN Z916=2;
ELSE IF Y916=1 & X16=0 THEN Z916=3;

IF Y917=. THEN Z917=.
IF Y917=0 & X21=0 THEN Z917=0;
ELSE IF Y917=1 & X21=1 THEN Z917=1;
ELSE IF Y917=0 & X21=1 THEN Z917=2;
ELSE IF Y917=1 & X21=0 THEN Z917=3;

IF Y918=. THEN Z918=.
IF Y918=0 & X29=0 THEN Z918=0;
ELSE IF Y918=1 & X29=1 THEN Z918=1;
ELSE IF Y918=0 & X29=1 THEN Z918=2;
ELSE IF Y918=1 & X29=0 THEN Z918=3;

PROC SORT; BY REGION ETHX;
TITLE 0=WET/WET 1=DRY/DRY 2=REALLY
  WET/PREDICTED DRY 3=REALLY DRY;
TITLE2 THREE DIMENSIONAL MODEL LATIT
  LONGI FACT2 (ETHNICITY);
TITLE3 ETHNICITY IN 5 CATEGORIES LT -1,
  -1 TO 0, 0 TO 1, 1 TO 2, GT 2;
PROC FREQ; TABLES Z881 Z883 Z886 Z887 Z891

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
PROC FREQ; BY REGION; TABLES Z881 Z883 Z886 Z887 Z891 Z904 Z906 Z907 Z908 Z909 Z910 Z911 Z916 Z917 Z918;
PROC FREQ; BY REGION ETHX ; TABLES Z881 Z883 Z886 Z887 Z891 Z904 Z906 Z907 Z908 Z909 Z910 Z911 Z916 Z917 Z918;
Robert Paul Sechrist was born March 29, 1955 in Lewistown, Pennsylvania. The family moved to Downingtown, Pennsylvania in 1959 where Robert lived until 1973 when he graduated from high school. At the University of Pittsburgh, he graduated with majors in anthropology and geography in 1977. In 1980, he obtained a master's degree in geography from the State University of New York at Binghamton.

Since 1980 he has been pursuing his Ph.D. in geography at Louisiana State University. On August 14, 1982, he was married to Gail L. Schlundt. He became employed by the Center for Governmental Studies at Northern Illinois University, in August 1984, where he still works as head of their programming and data processing staff.
Candidate: Robert P. Sechrist

Major Field: Geography

Title of Dissertation: A Comparative Analysis of Multi-Dimensional Diffusion Models: The Diffusion of the Prohibition Movement in the United States of America, 1876-1919

Approved:

[Signatures]

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination: November 26, 1985