VOTER PARTICIPATION IN NINETEENTH CENTURY BRITISH PARLIAMENTARY ELECTIONS

by

R. Morris Coats

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

APPROVED:

T. Nicolaus Tideman Co-chairman

Arthur T. Denzau Co-chairman

Daniel Orr

Gordon Tullock

H. Geoffrey Brennan

May, 1983

Blacksburg, Virginia

VOTER PARTICIPATION IN NINETEENTH CENTURY
BRITISH PARLIAMENTARY ELECTIONS

by

R. Morris Coats

(ABSTRACT)

A rational voter will vote only when the expected benefits outweigh the costs of voting. The costs of voting include not only transportation costs and wages forgone, but also any bribe to abstain. The product of the probability of affecting the outcome and the utility difference between the alternatives (measured in dollars) is the expected benefit of voting. The probability of affecting the outcome is affected by the voter's estimate of the closeness of the election and the number of voters.

Bribery was quite common in nineteenth century British elections. Before the secret ballot was introduced, votes of an individual were public record, making it easy to monitor votes case by a paid voter. After the secret ballot was introduced, monitoring paid voters was difficult, but it was still easy to monitor paid abstainers.

This dissertation examines evidence from British elections from this period, testing the assertion that the secret ballot decreased voter turnout along with other hypotheses concerning the effects of the costs and expected benefits on turnout. Multi-seat districts (but a single election) during this period of British elections necessitate an extension of the calculus of voting to include these cases.

ACKNOWLEDGEMENTS

I would like to acknowledge and express my gratitude to my dissertation co-chairmen, Professors T. Nicolaus Tideman and Arthur Denzau, for their advice, inspiration, encouragement and patience. I would like to thank Professors Daniel Orr, Gordon Tullock, and H. Geoffrey Brennan for the help they gave as members of my committee. I would also like to thank Professors Melvin J. Hinich and Robert Mackay who were former members of my committee.

There are so many who gave help to me, I could not begin to name them all, nor repay them back for their support. Among them are the consultants at the VPI&SU Computer Services and librarians at the Carroll Newman Library at VPI&SU, whose names I never knew. Also,. among them were many friends who generously welcomed me into their homes during the many weekends I came to Blackburg while living elsewhere, though I must especially thank Phillip Taylor and Andrew Economopoulos for sharing their roof and table with me for more than three weeks while I completed this dissertation and defended it. The help of my colleagues at Marshall University, especially Dr. Wendell E. Sweetser, Jr., was generously offered and greatly appreciated. I would also like to thank Juanita Willard, who typed this manuscript and countless drafts, giving up weekends and nights to work with me, whose good cheer seemed endless.

Finally, I would like to dedicate this dissertation to my parents, whose generosity, understanding, encouragement and support made this work possible.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND HISTORICAL BACKGROUND

> It is idle to expect that a study of Victorian elections will confirm either the myth that in the good old days of the nineteenth century men gave their votes for principle rather than for the party or self-interest, whereas in the twentieth century men gave their votes chiefly from habit, or the more recent myth that Victorian elections were festivals of beer and bribery. It is equally idle to expect of such a study that statistical sophistication which "psephologists" have displayed in analyzing British elections since 1945.
>
> H.J. Hanham in Elections and Party Management, 1959

I. Introduction

Hanham to the contrary notwithstanding, this dissertation brings some statistical sophistication to the analysis of Victorian elections, while adding support for his assertion that self-interest played an important role in Victorian election behavior. Although true motives of men may be impossible to discover, the assumption of selfishness provides a reasonable initial hypothesis in the study of human behavior. This study analyzes turnout in Victorian elections in terms that are readily interepreted as self-interest. The central result of ths study is evidence for a proposition put forth by Converse[1] and Rusk[2] and refined by Cox and Kousser,[3] that the introduction of the Australian ballot* causes turnout to decrease, other things being equal.

In the last decade of the nineteenth century and the first two decades of the twentieth century, voter participation in elections in

*Double-underlined terms are defined in the glossary at the end of Chapter I.

the United States dropped substantially. This led Burnham to contend that the voters in the last century were more interested in and informed about politics than are their twentieth century counterparts.[4] Burnham asserted that the take-over of the Republican party by business interests, represented by the McKinley nomination in 1896, alienated once active lower class voters by removing an effective alternative to capitalist domination from the political scene.[5]

Converse and Rusk countered Burnham in 1974 by suggesting that the decline in turnout was due to the reaction of reformers to widespread corruption of the period.[6] The reformers in this period introduced the secret ballot and also insisted on strict registration rules.[7] Converse and Rusk argued that these institutional changes made voting more costly.[8] They argued further that before these institutions were introduced the vote totals were swollen by fraud.[9] In other words, what happened when the secret ballot came into being was that fraudulent votes were reduced and this decreased voter turnout. So one must be careful in interpreting any decline in a figure such as voter turnout.

As Cox and Kousser pointed out, the debate between Burnham, on the one hand, and Converse and Rusk, on the other, involved two themes that have long concerned political scientists.[10] The first theme was the role of institutions in affecting human action.[11] The second theme was the external validity of results based on a particular technique or a particular data set.[12]

Before the introduction of the Australian ballot, agents of candidates could either pay the undecided voters and the supporters of the opposition to cast ballots for their candidates, or they could pay the opposition's supporters to stay at home on election day.[13] However, as the New York Democtratic state chairman said in 1900, "under the new ballot law you cannot tell how a man votes when he goes into the booth, but if he stays at home you know that you have got the worth of your money."[14]

Cox and Kousser suggested that before the Ballot Act, fraud inflated turnout because the acceptance of a bribe will "add a positive term to the right-hand side of their voting calculus equations."[15] They also suggested that after enactment of the ballot act, the form of the fraud should change to payments for abstention.[16]

Cox and Kousser tried to test for this change in the form of fraud by looking at newspaper accounts of fraud in rural New York.[17] They classified the accounts as either being inflationary or deflationary on turnout and used a chi square test to determine if the differences before and after the enactment of the Ballot Act in New York occurred by chance.[18] They found that the pattern of accounts (predominatly of inflationary before the Ballot Act and predominatly deflation afterward) was unlikely to have occurred by chance.[19]

In the next chapter the literature concerning the calculus of voting will be reviewed. The theory of the calculus of voting will be extended in the third chapter, with Chapter 3 focusing on multi-candidate, multi-winner, multi-vote elections. Chapter 4 develops a multivariate regression model of turnout based on the calculus of voting. This model is used in Chapter 5 to test for changes in the structure of turnout expected to result from two electoral reforms--

the extension of the franchise in 1867 and the introduction of the secret ballot in 1872. Finally, results will be summarized and and conclusions will be made in Chapter 6.

There are several reasons for focusing on turnout. First, it provides a way of testing several hypotheses concerning the calculus of voting. Second, changes in turnout can cause changes in electoral outcomes in close elections which will lead to changes in policy. Finally, turnout has been used by political scientists as an indicator of voter interest,[20] which could be misleading if variables other than voter interest are affecting turnout.

II. Historical Background

Before proceeding to the next chapter, some background in the institutions of nineteenth century British elections may help in understanding the reforms and their effects. Bribery was commonplace in nineteenth century British politics.[21] Candidates largely dissipated the benefits of a seat in Parliament by competing for office by bribery and other costly means. Election or campaign reform could be seen as the erection of barriers to competition by bribery and other campaign expenditures. The qualifications for voting were complex, and they varied, not only among the countries of England, Wales, Scotland and Ireland, but also among the various types of districts--county, borough, and university districts.

In a recent book, British Political Finance, 1830-1980, Michael Pinto-Dunchinsky remarks that the main features of British elections from 1830 to 1883 were bribery and high costs of elections.[22] Buying a seat was a familiar route to Parliament.[23] It is well known among economists that their intellectual ancestor, David Ricardo, purchased his seat in Parliament in 1819. This route was not even the cause of a decline in respect among M.P.'s. As the Westminister Review summed it up:

> It is a painful truth that a wealthy man, known to have bribed, nay actually be convicted of bribery, is not the whit less respected by the majority of the House.... That a candidate spent £10,000 in the corruption of a borough will no more exclude him from the general society of the House of Commons, than a man of fashion would have been tabood in the age of Congreve, because he had laid out a similar sum to corrupt a friend's wife.[24]

Although bribery was seldom the only consideration in an election, it was, nonetheless, a familiar one. Illegal inducements were offered to and accepted by voters regularly in most boroughs.[25]

Bribery was not the only means candidates, their agents and others had to influence votes. Employers would sometimes direct the votes of their workers.[26] Landlords had influence over their tenants, customers over shopkeepers, clients over soliciters, and clergy over congregations.[27] This sort of influence was universal.[28] Less widespread, yet a still common means of influencing the outcome, was violence. The banners and processions during electoral contests often led to riots.[29]

Another aspect of the period from 1832 to 1867 worth noting was plural voting, or being able to vote in more than one election.[30] To vote in a borough, a voter would have to live in, or within seven miles of, the borough in which he wanted to vote.[31] Anyone wanting to vote on an occupation qualifiction in a county had to prove he was a resident.[32]

There was nothing to prevent a man from being able to vote in several counties and a few boroughs, other than the cost and the requirement that votes be cast in person (except for university seats, where a vote could be cast by mail).[33] Plural voting enabled candidates to create nominal qualifications to vote for non-residents (sometimes called faggot-making), which was another corrupt practice that was used to tilt elections in one's favor.[34]

Not all of the corruption was in the form of direct cash payments for votes. Treating, the funding by candidates of innkeepers to provide refreshments (liquor), cigars, meals, and rooms to voters, was an even more common practice.[35] Candidates were also expected to transport voters to the polls, even in small borough districts.[36] To have a chance of winning, they also had to employ electors as cab drivers, messengers, canvassers, clerks, agents, and pollwatchers.[37]

The high costs of elections in this period (in constant pounds the costs per voter were ten times higher in 1880 than in 1979[38]) were also a drain on the candidates' purses. The costs of building or renting places to take the poll, the expenses of the returning officer and his poll clerks, and the fees and gratuities for special services of the sheriff and his men were all the responsibility of the candidates.[39] Candidates were also expected to contribute annually to local charities and political clubs and pay for the registration of favorable voters and objections to voters likely to oppose them.[40]

It was not merely their desire to serve the public, nor their thirst for power that led candidates to spend great sums for the chance of gaining a seat in Parliament. They were able to secure jobs and sinecures for friends and relatives.[41] Decisions about docks, canals and railways were made in small, private bills committees, putting an M.P. in a position to make large amounts of money.[42]

These rents were largely dissipated by bribery and treating. Members of Parliament attempted to reduce bribery, treating, and other expenses by making those acts illegal and by passing a number of measures to control their expenses. Among the objectives of the First Reform Act, passed in 1832, was the reduction of corruption and election expenses.[43] An act passed in 1841 made it less likely that seats gained by bribery could be retained, by allowing evidence about instances of bribery to be taken before proof of agency was established.[44] The penalties for bribery were also stiffened in 1854.[45] To discourage bribery by overwhelming the candidates with electors they would have to bribe, the Second Reform Act was passed in 1867 which greatly extended the franchise, especially in the boroughs (by 145%).[46] The Second Reform Act also regulated the expenses of the returning officers.[47] In 1868, the Corrupt Practices Act was passed in an attempt to reduce the leniency of punishments and the expense of trying election petitions.

The year 1872 produced two laws to combat bribery. The first, recognizing the difficulty of having pure national elections when the local elections (in particular, elections for returning officer) were fought with corrupt means, made fraud in municipal elections just as illegal as fraud in the Parliamentary elections.[48] The Secret Ballot

Act was passed in 1872 to raise the cost of monitoring the votes of paid voters, making the practice of buying votes less certain in outcome.[49] Three years later the returning officers' expenses were further regulated.[50]

Finally, after the most costly election in British history in 1880, the Parliament passed the Corrupt and Illegal Practices (Prevention) Act in 1883.[51] This act transferred property rights in the election rents from the voters to the incumbents by effectively limiting the campaign expenses of candidates, banning expenses for refreshment and transportation of voters, forbidding others from incurring election expenses without the candidate's approval, and requiring strict disclosures.[52] The law reduced both the declared expenses and the illegal expenses because the disclosure rules reduced the scope for bribery.[53] The campaigning cost per vote fell drastically after the Act.[54]

Other reforms to reduce the costs of elections provided for more convenient and less costly polling. Although the distance to the polls was generally greater in the county districts than in the borough districts, it was limited somewhat. An act passed in 1828 required that the polling booths to be divided into compartments with at least one compartment for every 600 electors.[55] Further, the county Parliamentary districts were divided into polling districts by the First Reform Act so that a voter would have to go no farther than fifteen miles from the property which enabled him to vote.[56] This was to reduce the expense of transporting electors to the polls.

The First Reform Act also limited the number of days of polling to two in both the counties and the boroughs.[57] Before the act, counties polled for fifteen days. Polling was restricted to one day in the boroughs in 1835 and in the countries in 1853.[58] This reduced the expenses of the returning officers and their men, as well as the wages of the candidates' workers, and even the bribery expenses.[59]

There are several aspects of the representative system of nineteenth century Britain that need explanation. First, the Scottish, the Irish, the English and the Welsh had different systems of representation, with different qualifications for voting.[60] Scotland and Wales were similar in that they both had groupings of boroughs (burghs in Scotland), which would form a single district of boroughs to elect one or more M.P.'s.[61] Second, the qualifications within each country varied, depending on the type of electoral district: county, borough, or university.[62]

Some districts returned more than one member of Parliament. The number of seats per district was not closely related to the number of electors. The county districts usually had two seats, though some had three, and one, the Isle of Wight, had only one. Some counties were divided up into several Parliamentary districts as well. The number of seats in the borough districts varied from one to four, while the university district had either one or two seats.[63]

In the multi-seat districts before the Second Reform Act, the voters could cast as many votes as there were seats for that district but could not vote for one person more than once.[64] The Second Reform Act had a minority clause which gave the districts with three or more seats something akin to proportional representation, by allowing voters to cast one fewer votes than the number of seats in the district.[65]

In 1832 and 1868, growing industrial towns were enfranchised (new borough districts were created) and some smaller boroughs were enlarged by extending their boundaries. In those years there were some redistributions in the number of seats, increasing the number of seats in some districts, decreasing the number of seats in other districts. From time to time, some of the especially venal boroughs were disfranchised altogether as a punishment to corrupt electors, causing those voters to be disfranchised unless they could qualify in another district or meet their county qualifications.[66]

The qualifications for voting were really quite simple: one's name must appear on the official electoral roll to be able to vote.[67] The qualifications to have one's name on the electoral roll were another matter. The franchise was closely connected with property, for the British saw those with a significant amount of property as having a particularly strong interest in Britain's survival, as Britain's survival was their greatest assurance of continued entitlement to the use of their property.[68] Gash aptly detailed the English and Welsh qualifications to be on the electoral rolls between 1832 and 1867:

> The electoral qualifications sanctioned by the English Reform Acte of 1832 were even more diverse and complicated than under the unreformed system. By the act the existing <u>county franchise</u> of 40s. <u>freehold</u> by inheritance was not disturbed but limitations were imposed on freehold tenure for life. Unless seized at the time of the passing of the act, none were entitled to vote in respect of such tenements unless they were in <u>bona fide</u> occupation of them or unless they were possessed of them by marriage, devise, or promotion to office or

benefice, or unless the property was of the annual value of £10. The county franchise was further extended to the possessor of the £10 copyhold, the £10 leasehold of not less than sixty years, and the £50 leasehold of not less than twenty years, and to the tenant of lands or tenements paying not less than £50 per annum in rent. It is worth noticing that a freehold was not necessarily land. It could be anything in the nature of property or interest arising out of the land, such as rent, tithe, shares in rivers or canals, or market tolls.

In the boroughs the central qualification for the franchise was occupation, as owner or tenant of one landlord, of buildings of the annual value of £10. But the famous £10 householder clause was hedged around by a mass of conditions and subsidiary franchises. Occupation, to be a qualification, entailed residence for the preceding twelve months at least. All rates and taxes had to be paid and the property rated to the poor rate before a legal claim to vote could be entertained. Tenants who compounded for rates with their landlords were thus placed in a peculiar and difficult position. In addition the old pre-reform franchises were allowed to continue in force during the lifetime of their owners. Freeholders and burgage tenants with the right of voting previous to the passing of the act were to retain the vote if they registered themselves and continued in actual possession of their freehold and burgage tenements. Burgesses or freemen with a previous right of voting were to retain their vote on condition of continued residence in the city or borough, or within seven statute miles thereof, but no freemen created after 1831 could vote unless they were qualified by birth or servitude. All other persons with a right of voting previous to 1832 were to retain the vote as long as they remained duly qualififed, except in the boroughs of Schedule A which were totally disfranchised. It was obviously the poorer classes that benefited in the main from these concessions as the rich were enfranchised in any case. The result was that the electorate was still wide and heterogeneous for many years after 1832 and only became narrower and socially more unified as the old franchise died out. But this was a process that required some twenty or thirty years for its substantial completion.

Many persons were of course qualified both for the old and the new franchises. Many persons too had qualifications for more than one constituency. On the other hand persons otherwise qualified were in some cases legally debarred from exercising the franchise. Receipt of parochial relief of alms was a disqualification in cities and boroughs though not in counties; no peers (except Irish peers), magistrates or officers were allowed to vote; aliens, women, minors, lunatics (except during lucid intervals), were also disqualified. Above all, for votes to be valid, all holders of the franchise had to be properly registered....

A fee had to be paid for registration. In the counties every elector had to pay 1s. at the time of claiming his vote; in the boroughs the elector paid 1s. at the time of registration and 1s. a year thereafter.[69]

In some boroughs, called nomination or pocket boroughs, the qualifications before 1832 was given to anyone who paid poor rates, and in others to every male resident for the last six months, not a charge on the poor rates (not receiving relief), and who had a family.[70] Although these boroughs had low qualifications, the elections were largely controlled by one man. Voters registered under these qualifications remained eligible after the First Reform Act as long as the borough in which they were registered was not disenfranchised.[71]

The Scottish Reform Act in 1832, established the £10 household as the qualification in the borough,[72] while in the counties, actual ownership of "land, houses, few duties, or other heritable subjects to the annual value of £10,"[73] qualified one to register. The franchise was also available to male

"leaseholders of property worth £10 annually on a life lease or one not less than fifty-seven years, leaseholders of property worth £50 on a lease of not less than nineteen years, and *bona fide* tenants of property worth £50... together with the existing owners of 'parchment votes' (i.e., *superiorities*) for the remainder of their lives."[74]

The Second Reform Act (1867), increased the electorate by 82.5 percent.[75] In England and Wales the borough franchise was extended to all householders who had actually paid their rates and had resided in the same borough for a year.[76] The county franchise was extended down to £12 rental value and to £5 leaseholders and copyholders.[77] The qualifications were the same for Scotland except there was no need for a "lodger clause," since tenants already had the franchise.[78]

At first glance, it is difficult to understand why Disraeli's Conservative government would support such a democratizing proposal as extending the franchise. However, a closer examination reveals a lopsided growth in the borough and county constituencies. The county districts were increased by 45 percent, the boroughs by 134 percent.[79] We will offer one possible explanation for the extension.

The central party organizations were growing in the 1860's, becoming more influential, and helping some candidates with election expenses.[80] A strategy that the governing party might take to strengthen its position would be to get the opposition to spend greater sums in places where their chances for winning was already great, diverting resources from the more marginal districts. The relationship between electorate size and election expenditure was known to be a positive one,[81] so it might have made sense for the Conservatives to try to enlarge the electorate in districts that tended to be Liberal strongholds by more than the strongly Conservative districts were increased. At the same time we would expect the Conservatives to try to strengthen their hold on their districts and weaken the Liberal hold on the Liberal districts. The English county constituencies were the strongholds of the Conservative party, the large boroughs the strength of the Liberals. This contention is supported by Hanham, who writes:

> The reformers in 1867 deliberately went out of their way to give the boroughs a franchise entirely different from that

> of the counties, so increasing rather than playing down the traditional difference between them. They came therefore to differ not only in size, in tradition, and in local government, but also in the whole way in which elections were fought. The counties became the strongholds of the old order, the boroughs, or rather bigger boroughs, became the field for experiment in 'democratic' political organization.[82]

The qualifications for candidates for Parliament should also be mentioned. For England and Wales these were property qualifications of a landed estate of £600 per year for county members and £300 per year for borough members.[83] In 1838 the qualifications could be met by personal as well as real property.[84] Scotland had no such property requirements for their candidates.[85] These qualifications were seldom enforced, however.[86] There were no residency requirements allowing a candidate to run for whatever seat he thought he could most easily (with least cost) win. There were even some candidates that ran in more than one district in the same election and won more than one race.[87] They could not hold more than one seat, though.

This gives rise to a puzzling situation. Although candidates could go to what they thought was the cheapest seat, county seats were consistently more costly than borough seats.[88] This raises the question of why some of the candidates contesting the county seats did not try to run for a borough seat, eliminating all profit opportunities. A reference giving election returns for this period supplies the reason for the persistent cost difference. Borough seats were more frequently contested, so a borough seat was not as valuable as a county seat.[89]

FOOTNOTES

[1]Philip E. Converse, "Comment on Burnham's 'Theory and Voting Research'," American Political Science Review, Vol. 68, 1974, pp. 1024-1027.

[2]Jerrold G. Rusk, "Comment: The American Electoral Universe," American Political Science Review, Vol. 68, 1974, pp. 1041-1042.

[3]Gary W. Cox and J. Morgan Kousser, "Turnout and Rural Corrpution: New York as Test Case," American Journal of Political Science, Vol. 25, No. 4, Nov. 1980, pp. 646-664.

[4]Walter Dean Burnham, "Theory of Voting Research: Some Reflections on Converve's 'Change in the American Electorate'," American Political Science Review, Vol. 68, 1974, pp. 1002-1023.

[5]Walter Dean Burnham, "The Changing Shape of the American Political Universe," American Political Science Review, Vol. 59, No. 1, pp. 7-28.

[6]Op. cit., Converse, pp. 1024-1027, see op. cit., Rusk, pp. 1028-1046.

[7]Loc. cit.

[8]Loc. cit.

[9]Loc. cit.

[10]Op. cit., Cox and Kousser, p. 646.

[11]Loc. cit.

[12]Loc. cit.

[13]Gary W. Cox and J. Morgan Kousser, "Turnout and Rural Corruption: New York as a Test Case," Social Science Working Paper 292, Pasadena: California Institute of Technology, October, 1979, pp. 9-14. The revised form of this paper has been cited in Footnote No. 3, above.

[14]Ibid., p. 14.

[15]Ibid., p. 11.

[16]Ibid., pp. 12-13.

[17]Ibid., pp. 13-17.

[18]Ibid., p. 14.

[19]Ibid., pp. 14-17.

[20]V.O. Key compares turnout in Virginia with turnout in Mississippi, and suggests that Mississippi's higher turnout means that Mississippi voters were more interested in public affairs--but Mississippi's elections were also much closer. Southern Politics In State and Nation, Alfred A. Knopf, New York, 1950, pp. 20, 28 and 246-253.

[21]William B. Gwyn, Democracy and the Cost of Politics in Britain, The Althalone Press, London, 1962, pp. 61-92. Norman Gash, Politics in the Age of Peel, Harvester Press, Limited, Hassocks, Sussex, 1977, pp. 154-172. Trevor Lloyd, The General Election of 1880, Oxford University Press, London, 1968, pp. 109-133. Michael Pinto-Duchinsky, British Political Finance: 1830-1980, American Enterprise Institute, Washington, D.C., 1981, p. 15.

[22]Op. cit., Pinto-Dushinsky, p. 15. Op. cit. Gwyn, p. 71-73.

[23]Op. cit., Gash, p. 154.

[24]Op. cit. Gwyn, p. 72, quoted from Westminster Review, 78 (1862), p. 36.

[25]Op. cit., Gash, p. 154.

[26]Op. cit., Gash, p. 137; op. cit., Lloyd, p. 123

[27]Op. cit., Gash, p. 175; Op. cit., Lloyd, p. 123.

[28]Loc. cit.

[29]Op. cit., Gash, 141. Gash devotes one chapter to bribery, one to violence, and one to influence and control, as means of competing for office.

[30]Op. cit., Lloyd, p. 123.

[31]Ibid., p. 124

[32]Loc. cit.

[33]Loc. cit.

[34]Ibid., pp. 124-125.

[35]Op. cit., Pinto-Duchinsky; p. 16, Op. cit., Gash, pp. 119-121.

[36]Op. cit., Gash, p. 19, see quote on page 108 of the present study.

[37]Op. cit., Pinto-Duchinsky, p. 16.
[38]Ibid., p. 27.

[39]Op. cit., Gywn, p. 22.

[40]Op. cit., Pinto-Duchinsky, pp. 16-17

[41]Ibid., p. 18.

[42]Ibid., p. 27.

[43]Op. cit., Gash, pp. 12-14.

[44]Op. cit., Gwyn, pp. 81-82.

[45]Ibid., pp. 83-84.

[46]Op. cit., Pinto-Duchinsky, p. 24.

[47]Op. cit., Gwyn, p. 23.

[48]Op. cit., Pinto-Duchinsky, p. 24.

[49]Op. cit., Gwyn, p. 89.

[50]Op. cit., Gwyn, pp. 25-26.

[51]Op. cit., Pinto-Duchinsky, p. 26.

[52]Loc. cit.

[53]Loc. cit.

[54]Ibid., p. 27.

[55]Op. cit., Gwyn, p. 43.

[56]Loc. cit.

[57]Loc. cit., and Op. cit., Craig, pp. xv-xvi

[58]Op. cit., Craig, pp. xvi.

[59]Op. cit., Gwyn, p. 47.

[60]Op. cit., Gash, pp. 3-64.

[61]Op. cit., Craig, pp. 497-563

[62]Op. cit., Gash, pp. 65-101. Since Ireland has been excluded from the present study because of lack of data, a discussion of Irish institutions will not be undertaken here.

[63]F.W.S. Craig, ed., British Parliamentary Election Results 1832-1885, 1977, MacMillan Press, London, for the data on seats and electorate sizes, passim.

[64]Op. cit., Craig, p xvi-p xvii.

[65]F.B. Smith, The Making of the Second Reform Bill, Cambridge at the University Press, London, 1966, pp. 212-213.

[66]Op. cit., Gash, pp. 95-96.

[67]Op. cit., Gash, p. 87.

[68]Op. cit., Gash, p. 19.

[69]Op. cit., Gash, pp. 86-88.

[70]Michael Brock, The Great Reform Act, Hutchinson and Co., Ltd., London, 1973, p. 18.

[71]Op. cit., Brock, p. 139

[72]Op. cit., Gash, p. 43.

[73]Op. cit., Gash, p. 42.

[74]Loc. cit.

[75]Op. cit., Smith, F.B., p. 237.

[76]Maurice Cowling, 1867, Disraeli, Gladstone and Revolution, Cambridge at the University Press, London, 1967, p. 14.

[77]Op. cit., Cowling, p. 14.

[78]Op. cit., F.B. Smith, p. 26.

[79]Op. cit., Gwyn, p. 33.

[80]Op. cit., Gwyn, p. 107

[81]Op. cit., Gwyn, pp. 34-35.

[82]H.J. Hanham, *Elections and Party Management*, Harvester Press, Hassock, Sussex, 1978, p. xxv.

[83]Op. cit., Gash, 105.

[84]Loc. cit.

[85]Loc. cit.

[86]Loc. cit.

[87]Op. cit., Craig, pp. xv.

[88]Michael Brock, *The Great Reform Act*, Hutchinson & Co., Ltd., London, 1973, p. 18.

[89]Op. cit., Craig, passim.

GLOSSARY

AUSTRALIAN BALLOT: A ballot, which, by law, is to be marked anonyomously.

*BURGAGE: A tenure by which real property was held in English boroughs.

*BURGESS: A citizen of a British borough.

COMPOUNDING FOR RATES: The combining with the landlord by the tenents, to pay the poor rates (paid as part of the rent).

COUNTY FRANCHISE: The right to vote in county Parliamentary districts.

*COPYHOLD: (1) a tenure of land in Enland and Ireland largely abolished by the Copyhold Act of 1894 by copy of court roll, at the will of the Lord, and according to the custom of the manor of which the land was a part; (2) an estate held by copyhold.

*FEE: (1) an estate in land held in feudal law from a lord on condition of homage and service; (2) a piece of land so held; (3) an inherited or hertiable estate in land.

*FEE SIMPLE: A freehold estate of inheritance in land or hereditaments that may last forever and may be inherited by all classes of both lineal and collateral heirs of an individual owner or grantee.

*FEE TAIL: An estate or fee granted to a person and his issue or a designated class of his issue that is subject to the possibility of reversion if there is no such issue or no alternative gift to a designated person in case there is no such issue. English Statute DeDonis of 1285.

*FREEHOLD: (1) a tenure of real property by which an estate of inheritance in fee simple or fee tail or for life is held; (2) an estate held by such tenure; (3) a tenure of an office or dignity similar to a freehold.

*LEASEHOLD: (1) a tenure by lease; (2) land held by lease; land held as personally under a lease for years.

OCCUPATIONAL QUALIFICATION: A qualification to vote by the occupation of certain real property.

*PEER: A member of British nobility (a duke, marquises, earl, viscount, or baron).

*POOR RATE: An assessment levied for relief of the poor.

PROOF OF AGENCY: Proof that one was acting on behalf or under the direction of another (a principal) for whom he works which makes the principal responsible for the actions of his agent.

RETURNING OFFICER: An election official who certified that a particular candidate (or set of candidates) were duly elected and entitled to be seated in Parliament.

*SCOT AND LOT: A parish assessment formerly laid on subjects in Great Britain according to their ability to pay.

*SINECURE: A job or position requiring little work and providing an income.

†SUPERIORITIES: In Scotland, real property and the franchise granted from that property were legally seperable, so that a landowner could sell the land and still retain the right to vote. These property rights, called superiorities, were then bought and sold, as were other property rights.

*<u>Webster's Third International Dictionary</u>, editor-in-chief, Phillip Babcock Gove, Springfield, Mass.: G. & C. Merriman, 1961.

†Op. cit., Gash, p. 42.

CHAPTER 2

SURVEY OF THE LITERATURE

I. Introduction

The first of the two lines of literature to be reviewed is the theoretical work on rational choice models of electoral participation. In this line of literature we shall first look at the development of the calculus of voting. We then look at the theoretical contributions which focused on specific hypotheses of the calculus of voting. We then review of three rational abstention theories in competition with the expected utility calculus of voting. We shall end our review of this first line of the literature work with a look at a work concerning optimal voter strategies under different election system.

The second line of the literature to be reviewed is concerned with empirical studies of rational absention. We shall first turn our attention to the more general empirical tests and then to tests of more specific hypotheses, and finally we shall survey tests of a competing theory of rational abstention, minimax regret, against the expected utility theory.

> *The White Rabbit put on his spectacles, "Where shall I begin your Majesty?" he asked.*
>
> *"Begin at the beginning," the king said very gravely, "and go on till you come to the end; then stop."*[1]

II. The Development of the Theory of the Calculus of Voting--First Line

A. The General Models

Economists in the 1950's had a strong interest in the government and its impact on the market and private choice, but little had been done in the study of democratic decision-making. There was little more than wishful thinking concerning governmental decisions. Benevolent despots were implicitly assumed. Downs' work, The Economic Theory of Democracy, was among the first to describe public sector decision-making with a rule paralleling profit maximization in the private sector.[2]

The theory of the calculus of voting is set forth in Downs' Chapters 3, 13 and 14.[3] The remainder of the book is concerned with other aspects of defining a decision rule for governmental behavior instead of the vote/abstain decision and will not be discussed here.

In the third chapter, Downs develops the idea of the party differential for a voter in a two-party system.[4] The party differential is the difference between the utility that the voter expects to gain from his favorite candidate's victory and the utility he expects from the victory of the other candidate.[5] These expectations, Downs explains, are extrapolations based, in part, on past actions and policies of the candidates and their respective parties.[6]

The probability that the voter will affect the election outcome is not discussed until chapter 13, which is concerned with the returns to information.[7] There, Downs points out that a citizen does not behave as if his vote alone determines the outcome of the election.[8] A citizen must discount the value of his vote by the probability that his vote will be decisive.[9] This discount results in the citizen's vote value. Downs produces this analysis to show that the voter has little incentive to gain political information.[10]

In his fourteenth chapter, Downs discusses the rationality of abstention.[11] He first considers the case of costless voting and concludes that any non-indifferent citizen will vote.[12] Although his vote value may be small, it will still outweigh the zero cost of voting.[13]

Downs also includes the value of voting per se, the value of voting just to vote.[14] He explains voting as having value for only one reason: if no one were to vote, there would be no government selected, democracy would collapse.[15] With no voting costs, those with preferences vote and democracy is saved. Indifferent citizens need not vote.[16]

When voting is costly, the voting costs may outweigh all gains to voting and nonindifferent citizens may abstain.[17] Downs argues that this could lead some indifferent citizens to vote.[18] For an indifferent citizen to vote, there must be higher voting costs for some citizens with strict preference than for some indifferent citizens.[19]

Downs allows for what he terms "irrationality" in the model--voting for social pressure reasons.[20] One can, of course, treat social pressure as a rational argument in a citizen's voting calculus. Downs points out that this leads to the sterile conclusion that all behavior is rational, for each act is a means to an end that could be valued.[21] Of course, if we were to take one step back and introduce some theory about the situations in which the social pressures to vote are more prounounced, if we try to explain this behavior while avoiding obvious tautologies, then the theory can be strengthened.

In the next work in this survey, Towards a Mathematics of Politics, Tullock views social pressures to vote as negative costs (benefits) that can often outweigh the costs of voting.[22] Another early contribution to the field, Riker and Ordeshook's "The Theory of The Calculus of Voting," takes the route of making Downs's "irrational reasons" a separate term of the formal model.[23]

Tullock is concerned with explaining why people do not become informed about political affairs, not whether they vote.[24] Following Downs, Tullock shows that a citizen's incentive to become informed about political matters is dependent on the effect the information has in altering his judgment about the benefits he would get from having his more favored candidate in office rather than the opposing candidate.[25] Greater information, according to both Downs and Tullock, increases the benefits to the citizen of having his favorite candidate win by increasing certainty about his judgment.[26] However, as Tollison and Willet point out in a later article, greater information can cause either an increase or a decrease in the expected utility differences.[27]

Although Tullock acknowledges that he is following the trail blazed by Downs, he makes some significant improvements to the theory of the calculus of voting.[28] First, Tullock expresses the decision to vote in a simple equation, $P_i = B_i D_i A_i - C_i$, where:

P_i = the payoff of voting to citizen i,

B_i = benefit expected to be derived from the success of citizen i's favorite candidate or party,

D_i = likelihood that the citizen i's vote will make a difference,

A_i = discount for citizen i's estimate of the accuracy of his judgment $0 < A_i < 1$, and

C_i = the cost of voting to citizen i.[29]

Downs discusses the decision to vote or abstain as if it were expressed in this equation form (although he has no "A" term), yet, as far as we can tell, Tullock was the first to express the decision in such a simple form. This basic form is followed in most discussions of the vote/abstain decision.[31]

Second, Tullock points out that any incentive to vote in order to provide a public good is likely to be quite small.[32] Tullock avoids use of Downs' per se value of voting, possibly because he believes the value of providing a public good vanishes as the electorate size becomes large.[33]

Third, and most important, Tullock shows that even if the citizen votes because social pressures to vote outweigh the costs of voting, there is still no significant incentive to become adequately informed.[34] The citizen is best off casting an uninformed vote.[35] What Tullock neglects on this point is that there may be social pressures to become informed, just as there are social pressures to vote. Some citizens may become informed about political matters just as do many about soap operas or some celebrity's life, through this type of information is likely to be superficial.

As mentioned above, another major early contribution to the theory of rational abstension is Riker and Ordeshook's article "The Theory of the Calculus of Voting."[36] Riker and Ordeshook begin with the decision to vote expressed in a Tullock-type equation:

$$R = (BP) - C$$

where: R is the payoff to voting expressed in utility;

B is the difference in utility between having the citizens favored candidate win and having the more other candidate win;

P is the probability that the citizen will cast the decisive vote; and

C is the cost to the citizen of the act of voting.[37]

To make the model more general, they allow for another benefit to voting. They suggest that there may be some benefits to the act of voting apart from those that come from altering the outcome.[38] Tullock's social pressure and special interest vote value and Downs' _per se_ voting value fit into this category.[39] Riker and Ordeshook give other reasons:

1. the satisfaction of voting to comply with a felt duty to vote;

2. the satisfaction of stating an allegiance to democracy;

3. the satisfaction of stating an allegiance to a political party or ideology;

4. the satisfaction of the act of voting — the enjoyment factor (voting as cheering in a spectator sport); and

5. the satisfaction from expressing one's belief in the effectiveness of stating one's satisfaction with the political system.[40]

Riker and Ordeshook point out that these are surely some of the reasons that people vote.[41] They define a term, "D", as expressing the satisfaction of voting that is not associated with altering the outcome.[42]

Another factor affecting the decision to vote is the expected cost of voting for one's favorite candidate.[43] This factor will enter only in elections involving open ballot rulès, where people can make good on threats.[44] The intrinsic value of voting, as Tullock pointed out earlier, merely changes the sign of the cost term.[45] The difference is that the intrinsic value of voting term seems to be pervasive, while the expected cost term exists only when others can tell whom someone else voted for.[46] In modern elections, this seldom occurs (except for legislative roll-call votes). In the elections under consideration in this study this term may be important. Intimidation could either decrease or increase turnout. Intimidation could increase participation by reversing the signs on the utility difference term, making the initially less-preferred candidates the more preferred and making this utility difference greater than it was initially.

As their final model, Riker and Ordeshook have a description of the vote or abstain decision as determined by the equation R = PB - C + D.[47] They note that although the C and D terms are by far the most important, the expected benefit term, PB, can be important, making the difference between voting and abstaining for those citizens for whom the C and D terms are about equal.[48] Whether one's vote is decisive is a function of not just the number of voters, but also the citizen's expectation of the number of voters favoring each candidate.[49] Using a continuous subjective probability function as an approximation of the discrete probability of candidate 1 receiving X votes, and holding constant the total number of votes to be cast by others than decision maker himself, V, Riker and Ordeshook show the probability that one's vote will be decisive as r in Figure II.1.[50] Notice that candidate 2's probability

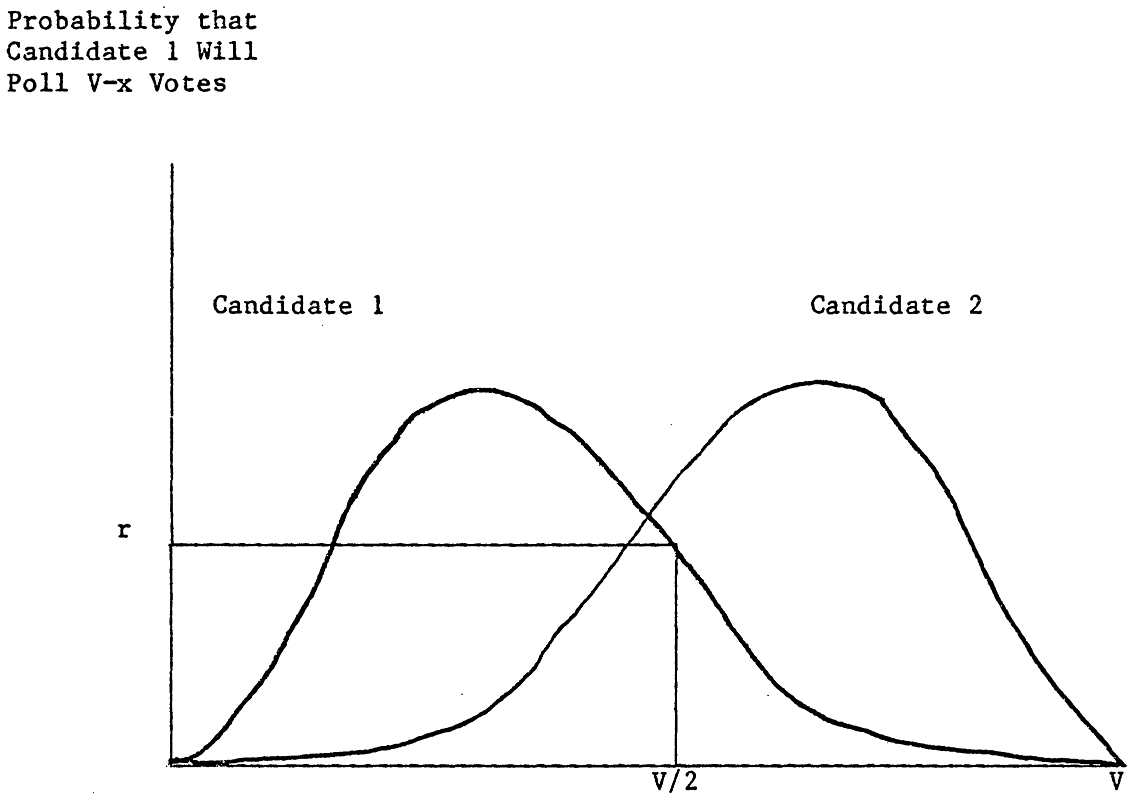

FIGURE II.1

THE PROBABILITY OF PARTICULAR OUTCOMES FOR CANDIDATES 1 AND 2 of receiving x votes must be candidate 1's probability of receiving V-x votes. Notice also that the assumption of a constant number of votes cast, excluding the vote of the decision maker himself, is not consistent with the nature of the model. If each citizen is able to make this vote/abstain decision, then the votes cast could be anywhere between 0 and the number of eligible voters. In Chapter 4 we relax their inconsistent assumption, showing it to be unnecessary.

One result of their model is to show that expected closeness is an important determinant of probability of affecting the outcome.[51] Thus, the probability is not determined simply by the number of voters. In the next two chapters, we show the relationship of the probability of altering the outcome with the expected closeness, the number of voters, and the degree of certainty about the expected closeness, using a multinominal probability function.

Riker and Ordeshook's empirical work in "A Theory of the Calculus of Voting," will be discussed in detail in the next section. It is, as far as we know, the first empirical test of the Downs hypothesis. It is sufficient to say here that the theory passed this first test.

The next major work in the development of the theory of the calculus of voting is the one undertaken by McKelvey and Ordeshook.[52] They generalize the theory of the calculus of voting to deal with multi-candidate races.[53] One unsurprising conclusion is that a citizen may vote for a candidate that is not his most preferred if the probability of this candidate winning is quite remote.[54] Chapter 3 is built directly on their work, extending the model to deal with multi-candidate, multi-winner elections. Their work is so vital to ours that a full discussion of it is undertaken is in Chapter 3.

McKelvey and Ordeshook extended the calculus of voting to multicandidate elections.[55] Some elections involve not only three or more candidates, but two or more winners.[56] Real elections, such as those for cheerleaders and for city councils, sometimes involve multiple winners.[57] In our data set we have some multi-winner observations. The question unanswered by the literature is "does having more than one winner in a contest seriously alter the effect of closeness on turnout?" We answer this question in the next chapter before we proceed to empirical tests that include observations with multiple winners.

B. Specific Models

There are several other works that are of a more specific nature, works that focus on specific terms in the voting calculus equation. Tollison and Willett[58] and Zechman[59] focus on the utility of voting due to changing the outcome of the election (Downs' party-differential term[60]). Beck,[61] Margolis,[62] Good and Mayer,[63] Ledyard,[64] Palfrey and Rosenthal,[65] Coats,[66] and Chamberlain and Rothchild,[67] add to the theory by showing how the objective probability of a tied vote is computed. Finally, Frey,[68] Fraser,[69] Russell,[70] and Niemi[71] discuss the costs of voting and the costs of non-voting, that is, Riker and Ordeshook's "C and D" terms.

1. The Party Differential (B)

The party difference term, B, is basically the utility difference between the two parties (or candidates or alternatives). But what things determine this utility difference? Is the effect of information the same as suggested by Downs[72] and Tullock?[73] Tollison and Willett[74] answer the latter question, Zechman[75] the former.

Tollison and Willett[70] correct an error made by both Downs[77] and Tullock,[78] who asserted that increased information will unambigously lead to a greater "B" term. Tollison and Willett show that greater information can either increase or decrease the expected value of the party differential by moving the distributions of utility gains from the two parties closer together or by moving them farther apart.[79] This means that greater information may have no systematic effect on turnout. Tollison and Willett also point out that Downs' per se vote value involves a prisoners' dilemma problem and will result in a cooperative solution after repeated plays only in a small-numbers context.[80] This, however, seldom occurs in a modern democracy. This illustrates the need to accept, as part of the model, Riker and Ordeshook's "D" term, (i.e., the utility of voting for reasons other than to change the winner).[81]

Zechman approaches the party differential term from a slightly different angle.[82] He appears to reinvent the wheel in his 1979 paper in Public Choice by giving arguments that are similar to those given in Downs' Chapter 3.[83] What Zechman does contribute, however, is more precise modelling of Downs' party trend and a Bayesian explanation of the effects of new information on the voter's expectations of future party programs.[84] This gives incentives to party managers to stress different aspects of their party's record,[85] a result similar to one by

Aranson, Hinich, and Ordeshook in "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives."[86]

What conclusions concerning the party differential can be drawn from these two articles? The first conclusion is simply that increased information, by itself, has an ambiguous effect on the size of the party differentials. Since uncertainty concerning the future policies of the parties will lead voters to rely on expectations of these values, the voters' expectations will be based, primarily, on the parties' past policies. These may not be known to all, which leads party managers to place different emphases on different parts of their records, which is the second conclusion.

2. The Probability of Altering the Outcome (P)*

A key element in much of the rational voter participation models is the probability of affecting the ordering of candidates. Beck calculates the probability of the reference voter breaking a tie in a two candidate election, given various parameters for the binomial distribution.[87] He observes a distinct "knife-edge" where the probability of breaking a tie falls very rapidly as the probability of voting for the first candidate moves slightly away from 0.5.[88]

Good and Mayer develop a Bayesian model of a voter's probability calculations.[89] They use a Beta distribution to represent a voter's prior subjective probability distribution because of the flexibility of the Beta in representing various prior distributions.[90] They show how

---

*This section is taken verbatim from a paper "The Probability of a Tied Election...One More Time," presented by R. Morris Coats and Wendell E. Sweetser at the 1983 Public Choice Meetings in Savannah, Georgia.

a voter would revise his priors with binominal sampling.[91] They find the distribution of the probability of casting a decisive vote to be flatter than Beck's distribution.[92]

Margolis extends Beck's analysis by introducing uncertainty about the binomial parameters.[93] Margolis uses the normal distribution as an approximation to the binomial and shows that uncertainty about the parameters reduces the probability of any particular voter casting a deciding vote when the election is expected to be close (P = .5) and increases the probability when the election is not expected to be close (e.g., P = .55).[94] The knife-edge can be dulled under some circumstances.

Finally, Chamerlain and Rothschild rigorously reproduced the knife-edge result using a mixed binomial distribution in modeling the uncertainty of the parameters.[95]

It would seem that little more need be said about the probability of a tie in a two-candidate election. However, with the exception of the Good and Mayer model, the voter making the probability calculation does not consider the possibility of other voters making the same kind of calculation as he is and possibly abstaining.[96] In all of the calculations of the probability of a decisive vote, the number of voters other than our one decision maker, is known and fixed.

Good and Mayer mention abstention but do no calculations that allow for abstention.[97] They claim that the expected abstention rate would need to be highly variable to affect their probability calculations.[98]

Coats[99] and Palfrey and Rosenthal[100] show that the probability of casting a decisive vote should be the sum of multinomial probabilities over all possible tied outcomes, yet they do no calculations. Ledyard[101] and Palfrey and Rosenthal[102] show that in the rational voter's decision calculus, the probability of casting a decisive vote and the abstention rate must be determined simultaneously.

3. The Costs of Voting (C) and the per se value of voting (D)

The costs of voting and the intrinsic benefits of voting are by far the most important variables in magnitude in the voting decision, because the probability of altering the outcome of most elections with a single vote is so small. There is some debate concerning the factors involved in the cost of voting.[103] An important question is whether or not the wage rate should be used to calculate the value of a person's time spent voting. It could be argued that the use of the wage rate to identify the opportunity cost of voting is appropriate, not only when people take time out from work to vote, but also when they take time out of leisure activity. Therefore, other things equal, one would predict that higher income individuals will participate less in voting because of their higher opportunity costs.

However, high income individuals vote more often than the poor. In a short but provocative paper, Bruno S. Frey provides one solution to this paradox by showing that the "all things equal" assumption is violated.[104] High income individuals, he suggests, also tend to be more productive at political activities, it take them less time to become informed on the issues and to untangle the issues since they tend to have mental jobs instead of physical jobs.[105] This higher productivity in political activity leads to lower time lost from work (or leisure) to vote, which vitiates the wage rate effect on opportunity cost.[106] It is analogous to employing a skilled worker instead of an unskilled one.

Fraser,[107] Russell,[108] and Niemi[109] critize Frey[110] (Niemi also critizes Tollison and Willett on this point), suggesting that voting will have a lower opportunity cost if the time is taken away from leisure time. Their criticism is correct if there is something that prohibits voters from making marginal adjustments in the number of hours they work per week.[111] If the voter were able to make adjustments in his work time, then it makes no difference whether time is taken out from work activity or leisure activity.

The discussion of the costs and benefits of voting not associated with affecting the outcome of the election is advanced by Niemi in his Public Choice note.[112] His contribution to the discussion is his focus on the variation in the costs of voting across elections that affects each voter in the same direction (especially from high-stimulus to low-stimulus elections) which result in varying turnouts across elections.[113] He looks at several costs of voting: the cost of remembering, the costs of information gathering, and the costs of getting to the polls.[114] All of these are higher in low-stimulus local elections than in high-stimulus national elections.[115] The costs of not voting (the gains from voting not associated with affecting the outcome) are higher in high-stimulus election than in low-stimulus elections, he points out.[116] That is, embarrassment from admitting that you didn't vote is lower in low-stimulus elections than in high-stimulus elections simply because people more often tend to ask if you have voted in high-stimulus elections than in low-stimulus elections.[117]

The major question concerning the cost of voting is whether or not the wage rate gives us any information about the opportunity cost of voting. The answer we put forth is that it provides such information only if the worker can adjust marginally the number of hours he works.[118] A possible test of this proposition would be to compare proportions of voter turnout for self-employed with other employed individuals with similar wage rates.

C. Other Rational Choice Models of Voting

We will now consider three other rational choice models of voting. The first model is Ferejohn and Fiorina's minimax-regret model.[119] It uses Savage's decision rule[120] for making decisions under conditions of Knightian uncertainty. The second model is that of Strom,[121] who uses "regrets" from the Ferejohn and Fiorina analysis and the corresponding subjective probability of the occurrence of that state to develop an expected loss minimization model of voter turnout. The third model is Hinich's voting-as-political-contribution model.[122]

John Ferejohn and Morris Fiorina offer one possible rational choice alternative to the Downsian expected utility model of the vote/abstain choice, the minimax-regret criterion.[123] Savage's mimimax-regret criterion is similar to the more widely used minimax-loss (or maximin-utility) criterion.[124] The matrix of utility payoffs is transformed into a matrix of "regrets".[125] The minimax-regret decision-maker then chooses that action for which the greatest regret is smallest.[126]

Ferejohn and Fiorina's analysis led to five major results. First, in two-candidate races, the minimax-regret decision-maker will vote for his favorite candidate as long as his utility gain from having his candidate win is four times the utility loss of the voting act (costs of voting).[127] Second, in three-candidate races, a minimax-regret decision-maker will never cast a vote for any candidate other than his most preferred candidate.[128] Third, the voters following the minimax-regret criterion will vote, rather than abstain, under conditions that are weaker than the conditions for voting by an expected utility maximizer.[129] Fourth, the higher the cost of voting the more likely it is that abstention will be the minimax-regret strategy.[130] Last, the higher the utility from one's second choice the more likely it is that abstention will be chosen by a minimax-regret decision-maker.[131] Ferejohn and Fiorina are surprised at this last result, yet this is the same conclusion one finds in Downs' analysis.[132]

Mayer and Good argue that voting does not occur in an environment of Knightian uncertainy, but rather an environment where subjective probabilities can be formed by the potential voters.[133] Tullock criticizes minimax-regret by pointing out that almost no one votes for his favorite candidate--himself![134] Beck argues against mimimax-regret decision-making by the analogy that a minimax-regret decision-maker would forego the small certain gain of buying a newspaper if he had to risk crossing the street to do it.[135]

Strom suggests a model that is a hybrid of the Downsian expected-utility model and the minimax-regret model.[136] Essentially, the "utility" for each action/state pair is the negative of the regret element.[137] Each "utility" is then paired with the probability of that particular state of nature.[138] This model is difficult to defend because these "utilities" seem no better than those in the original utility matrix. If the individual is risk averse, it will show up in the shape of his utility function. Strom is also mistaken in suggesting that minimax-regret decision-making yields the same results as equiprobable expected utility models.[139] In the latter, the result will depend on how the states of nature (the various outcomes) are defined and differentiated.[140]

In another attempt to square rationality with the reality of widespread voting without postulating an all-encompassing "D" term, Hinich looks at voting as making a small contribution to your favorite candidate.[141] Like Strom, Hinich modifies the usual Downsian model and hypothesizes a gain for the voter if he votes for the winner or a loss if he votes for the loser.[142] Hinich's model is very close to Strom's model, but does not rely on any assumed risk aversion of the decision-makers.

What these models, particularly the minimax-regret model, contribute is additional rational explanations of abstention. The minimax-regret model, however, is only one possible model for decision-making under conditions where voters have no information concerning the closeness of the race on which to base subjective probabilities. This means that a test of this hypothesized model can not be based on whether or not the voter expects a close race, for this contradicts the basis of the model.

Although not a work concerning the calculus of voting directly, Brams and Fishburn's article, "Approval Voting" in the American Political Science Review is quite important to the present work.[143] From a very different approach, we get results which are very similar to theirs.

Under approval voting citizens mark the names of candidates of whom they "approve."[144] At most, a citizen will vote for all candidates except his least favorite, because voting for all candidates would just add one vote to each candidate's total, leaving the candidates' relative positions unchanged.[145] Approval voting can therefore be thought of as a non-weighted, multi-vote system of voting where the number of votes a citizen can cast is effectively limited to one less than the number of candidates.

In order to compare approval voting with other non-weighted single-ballot elections, Brams and Fishburn first prove several general propositions about non-weighted, single-ballot voting systems.[146] To explain their basic results, let us define some of their terms:

1. A strategy is a set of candidates a voter is considering voting for.[147]

2. A strategy is said to be feasible if voting for all of the candidates in the strategy is allowed in a particular voting system.[148]

3. A set of candidates, B, is said to be <u>high</u> for the voter's preference ordering iff whenever it contains any given candidate, a, it contains all candidates preferred to a.[149]

4. A set of candidates, B, is said to be <u>low</u> for the voter's preference ordering iff whenever it contains any given candidate, a, it contains all candidate which are less preferred than a.[150]

5. A strategy, S, is <u>admissible</u> iff it is feasible and there does not exist another feasible strategy, T, which dominates S.[151]

6. A strategy, S, is said to be <u>sincere</u> iff S is high for the voter's preference ordering.[152] A voting system, s, is said to be <u>sincere</u> if all admissible strategies for s and voter's preference orderings are sincere.[153]

7. A voting system is said to be <u>strategy proof</u> iff exactly one strategy is admissible for that voting system and a given preference ordering of the candidates.[154]

The first result of Brams and Fishburn is that a strategy, S dominates the strategy, T, iff S ≠ T, S intersects ~T is high for the voters preference ordering, T intersects ~S is low for the preference ordering, and neither S intersects ~T nor T intersects ~S is the set of all possible candidates.[155] The propositions we prove in the next chapter can be considered corollaries to this first result of Brams and Fishburn.[156] They show that although approval voting has many more possible strategies than many other voting systems, it does not give the voter many additional admissible alternatives.[157]

They also argue that there is reason to expect approval voting to be more frequently sincere, more frequently strategy-proof, and to produce a Condorcet winner more frequently than any other non-ranked, single-ballot voting system.[158]

Brams and Fishburn make one statement which seems to have little basis in positive theory. They state that since approval voting encourages sincere voting it would probably produce higher voter turnout, yet they give only flimsy reasons for this to occur.[159] From the theory of the calculus of voting it is easy to see that a voting system will encourage higher turnout if it expands a voter's choice set, the set of admissible strategies, other things being equal. Approval voting has the largest possible voter's choice set among non-weighted single-ballot voting systems, so this will work in the direction of increasing turnout.[160] Yet other things will not be equal, particularly the number of votes cast by others, which will increase the uncertainty concerning the probability of one voter's ballot affecting the outcome.

III. The Empirical Literature

We shall now show the necessary connection between the empirical models and the theoretical models we discussed in the previous section. We shall also relate our own empirical work to the works being reviewed.

First, we shall discuss the more general tests of the theory of the calculus of voting.[161] Second, we look at two articles that discuss the possible simultaneities among turnout, closeness, and campaign effort (expenditure). Finally, we review several tests of the minimax-regret model.

A. Tests of the General Models

We first examine Riker and Ordeshook's test of their own theory using an analysis of variance technique, after holding other variables constant.[162] Brody and Page also use analysis of variance to test two competing hypotheses about abstension.[163] The other four articles in this section, Barzel and Silberberg;[164] Tollison, Crain and Pautler;[165] Silberman and Durden;[166] Settle and Abrams[167] all use regression analysis in their tests. The last three of those four have their roots in the Barzel and Silberberg article, as does our own work since we use some of the same explanatory variables.

Riker and Ordeshook test their theory using data gathered by the Survey Research Center of the University of Michigan from survey questions asked during the 1952, 1956, and 1960 Presidential elections.[168] The response to these questions are used as proxies for Riker and Ordeshook's "P", "B" and "D" terms.[169] The proxy for the "P" term is the respondent's perception of the closeness of the race.[170] The "B" term is proxied by how much the respondent "cared" about the outcome of the election.[171] The "D" term is based on a sense-of-citizen-duty scale which the Survey Research Center constructed from responses to four questions about the citizen's beliefs concerning his duty to vote.[172]

Across respondents with any particular D (i.e., holding D constant) an observation is placed into one of four cells, one for P and B both high, one for a high P and a low B, etc.[173] For each cell, they compute the fraction of respondents who voted.[174] In ignoring C, they are assuming that C and D are not correlated and that C is relatively constant in each group.[175]

Their theory suggests that the proportion of respondents who voted in a particular election will be highest in the cell with high P and B terms and lowest in the cell with low P and B terms.[176] They also predict that the higher the D, the higher will be the proportion of respondents who voted.[177] They have only a few cells inconsistent with their hypotheses.[178] They computed the probability of having such a low number of invalidating cells as one in 20 million, supporting their theory.[179]

The next major test of the Downs theory of rational abstension is by Barzel and Silberberg.[180] Instead of using survey data, as Riker and Ordeshook do, Barzel and Silberberg use actual vote outcomes.[181] They use regression analysis to test their contentions about voter turnout and the probability of affecting the outcome.[182] Turnout in gubernatorial elections as a fraction of voting age population is the variable to be explained.[183] The first explanatory variable is the ratio of votes for the winner to the total vote, a proxy for the anticipated closeness of the election.[184] The total voting-age population is used as an explanatory variable because the probability of affecting the outcome also depends on electorate size.[185] As a proxy for intensity of feeling about the outcome, they use the percentage of the voting-age population who are registered voters.[186] They use the absolute difference in the percentage of Democratic votes between the current election and the previous election as a measure of the variance in the estimation of the election outcomes.[187] To measure the effects of other issues being on the ballot, they have a presidential dummy and a senatorial dummy.[188] As will become apparent, our work is similar in some respects to that of Barzel and Silberberg.

As expected, they find that the larger the winning majority, the lower the turnout, other things being equal.[189] Voting-age population and the difference in fraction of Democratic vote percentage also have a negative impact, as anticipated.[190] Again, as expected, the two other issue dummies have a significantly positive impact on voter turnout.[191] All of the variables are significant at the 10 percent level.[192] Their use of the ex post majority size and electorate size is similar to what we have done in the present study.

B. Tests of More Specific Hypotheses

1. The Effect of Information

Downs and Tullock both stressed the importance of information in the decision to vote.[193] Tollison and Willett cleared up the theoretical discussion by showing that the effect of greater information on the vote/abstain decision is ambiguous.[194] Tollison, along with Mark Crain and Paul Pautler, follow up Tollison and Willett's work with an empirical exploration of the effect of the information on voter turnout.[195]

They begin with Barzel and Silberberg's empirical model, modifying it by including three more variables, both separately and in combinations:

1. free time given to candidates on non-network broadcasts, and

2. total broadcasting expenditures for all offices divided by voting-age population.
3. An interaction term which is the product of the winner's percentage of the total vote (ex post) and total daily newspaper circulation expressed as a fraction of voting-age population.

Since they use data from 1970 gubernatorial elections, a presidential dummy is unnecessary.[196]

They found that the broadcast expenditures can not be said to affect total turnout.[197] They note, however, that political advertising may affect the "market shares" of the candidates.[198] The interaction term (closeness times newspaper circulation) and the free-time variable were quite significant and their inclusion raised the R^2 of the model from .50 to .68.[199]

Tollison, Crain and Pautler, although honest about their search among plausible models, do not seem to recognize that they have no tests of their contentions. They begin with a sample size of twenty-nine and have six parameters to estimate.[200] Then, they add variables in various combinations, entering the world of unknown sampling properties and invalid tests.[201]

Further, they point out a simultaneity problem with their suggestion concerning "market shares" of the candidates and relative broadcast expenditures, yet fail to follow up with the actual investigation.[202] Although Tollison, Crain and Pautler's work helps in further specifying an empirical model of turnout, their method of "testing" is precisely what we shall seek to avoid.[203]

Silberman and Durden identify eight factors that they associated with voter turnout based on the expected-utility hypothesis of voter behavior literature.[204] These factors are: (1) income; (2) percentage of population that is black; (3) percentage of population that is female; (4) percentage of population that is over 65 years of age; (5) the civilian unemployment rate; (6) the ratio of winner's votes to the total vote; (7) the voting age population; and (8) the presence of other issues on the ballot as proxied by dummy variables for senatorial and gubernatorial elections.[205]

Based on the work of Frey they hypothesize that voter participation should vary directly with income.[206] Since women and blacks have historically had little political influence on legislation and have been disenfranchised by both legal and illegal methods, they hypothesize that percentages of the population with these characteristics should be negatively related to turnout.[207] The over-65 age group is mostly composed of retired citizens and, therefore, of citizens with low opportunity costs of voting.[208] The citizens over 65 also tend to have strong feelings concerning "senior-citizen" issues and are seen as having large benefits from voting.[209] The civilian unemployment rate tests a similar hypothesis: that people with low opportunity costs and intense feelings about an issue tend to vote more.[210] The ratio of the winner's vote to the total vote and the size of the voting age population are both factors that affect the probability that a single vote will alter the outcome of the election.[211] Very similar measures are employed in the present work and in others surveyed here.

The dependent variable that Silberman and Durden use is the proportion of the population that cast votes in Congressional races in 1962 and in 1970.[212] They run regressions on each cross section (1962 and 1970) and a pooled cross section.[213] They found that all coefficients are of the proper sign and only the gubernatorial coefficient is insignificantly different from zero.[214] The 1962, 1970, and the pooled regressions have R^2's of .74, .54, and .60, respectively.[215] One interesting result is that income (median income) and race, although still important factors, had significantly smaller effects on voter turnout in 1970 than in 1962.[216] The remaining factors had relatively stable coefficients.[217]

The Civil Rights movement in the 1960's may be responsible for the decrease in the negative effect of percent black on turnout.[218] The findings of Silberman and Durden support a very important hypothesis of the expected utility theory of voter turnout--that the probability of a close election affects turnout.[219] Our own work follows Silberman and Durden in using both electorate size and majority size to capture the effect of the probability of a tied election on turnout.[220]

Settle and Abrams extend the empirical literature on rational voter participation.[221] Their empirical model is eclectic, borrowing from Riker and Ordeshook;[222] Tollison, Crain, and Pautler;[223] Silberman and Durden;[224] and Barzel and Silberberg.[225] They use per capita campaign spending, TV spending and radio spending to test for the information effect that Tollison, Crain, and Pautler proposed.[226] They include a majority size variable (<u>ex post</u>) much like that found in

Silberman and Durden and Barzel and Silberberg.[227] Following Silberman and Durden they also include an income term.[228] What is new to their work is the inclusion of a term for size of the "legislative profits" to be distributed.[229] They also include dummy variables for third-party candidate effects and the effect of the extension of the franchise to women in 1920.[230]

We are disturbed by one paragraph in Settle and Abram's paper, mentioning their experimentation in model building:

> In addition to these above mentioned variables, we experimented with a number of others that seem potentially important in explaining voter participation, but in practice have little effect on either the explanatory power of the regressions or the estimated coefficients of the other variables...[231]

One result is worth mentioning. They found income real GNP per capita to be negatively related to voter turnout, which they explain as being in keeping with the opportunity-cost-of-time argument, but in conflict with earlier studies because of model misspecification in the earlier works.[232] Another explanation may be that, since they use time-series data, there may happen to be a chance correlation between per capita real GNP growth and population growth or electorate size growth, i.e., that their model may be misspecified by not including voting-age population.

We find their work useful in pointing out the possible effect of franchise extension. In our own data set, British Parlimentary elections from 1860-1880, we find a large increase in the electorate size. In Chapter 4 we will use a Chow test to discover if there is any significant change in turnout or its relationship with independent variables that might be attributable to the increase in electorate size.

Political scientists have suggested the one important reason for abstention is voter alienation; this occurs when all choices faced by the voter have negative values.[233] The rational abstention literature suggests that a voter is as likely to vote when the choices have negative evaluations as when they have positive evaluations, as long as the differences between (or pairwise differences among) the alternatives are the same.[234] Brody and Page attempt to test these competing contentions by first cross-tabulating turnout by degree of alienation (evaluation of most favored candidate as either negative, neutral, slightly positive, or very positive) and then indifference and turnout (evaluation of the greatest difference in evaluation between any two candidates as either no difference, minimum difference, slight difference, moderate difference, or maximum difference).[235] Brody and Page find that both alienation and indifference separately contributed to abstension but that indifference is more strongly related to abstention (χ^2 for alienation cross-tabulation was over 30 and for indifference it was over 93).[236] A cross-tabulation between alienation and indifference would have been helpful, since any substantial correlation between them (which is quite possible) would make the test meaningless.

Brody and Page also find the correlation between a voter's most-favored candidate and the candidate he voted for stronger for the two major-party candidates than for third-party candidates (those viewed as having little chance to win).[237] This result supports McKelvey and Ordeshook's theory of multi-candidate elections.[238]

These articles point out some of the important determinants of turnout. These determinants are electorate size, presence of other issues, income, large extensions of the franchise, the size of the "pork barrel" (government expenditures), and closeness of the election. As we have seen, all of these works, with the exception of the last one (which does not make the test) support the basic contention that the closeness or expected closeness affects turnout. Our work also supports this contention. The next two articles discuss the causality in the relationship between turnout and closeness.

Two related empirical articles having an impact on our thinking are Seidle and Miller's "Turnout, Rational Abstension and Campaign Effort,"[239] and Denver and Hands' "Marginality and Turnout in British General Elections."[240] Seidle and Miller look at various causal relationships among marginality (closeness), campaign effort, and voter turnout.[241] One causal chain they examine is suggested by Denver and Hands: that elections that are expected to be close bring forth higher election expenditures, which hypothetically bring forth greater participation.[242] Although Seidle and Miller seem to suggest a possible simultaneity, they use OLS procedures.[243] In Chapter 6 we suggest that the Denver and Hands' hypothesis may be correct and can help explain some of the changes in the regression coefficients with the enactment of the Secret Ballot Act.

C. Tests of Minimax-Regret

Next, we turn to several articles that focus on testing two competing rational hypotheses of abstention: the expected utility hypothesis and the minimax-regret hypothesis. The key difference between the two is the importance of the probability term. In minimax regret, it plays no part in the decision, whereas in the expected utility framework it is the term by which utilities are discounted. Some analysts, including the major proponents of minimax regret, Ferejohn and Fiorina,[244] as well as Aldrich[245] and Cain,[246] have focused on the statistical importance of the probability term. Cain and Black[247] each point to another difference between the hypotheses: whether or not voters sometimes vote for their second choices. The minimax-regret hypothesis suggests that this will not occur.[248]

Ferejohn and Fiorina attempt to test their minimax-regret model of rational voter behavior against the Downsian expected utility maximization model.[249] They use data from the 1952-1964 presidential election surveys conducted by the University of Michigan Survey Research Center.[250] They employ Estimated Generalized Least Squares to estimate the binary response model $Y = \gamma + B_1X_1X_2 + B_2X_2 + \varepsilon$, where

$$Y = \begin{cases} 1 \text{ if respondent is a voter;} \\ 0 \text{ if respondent ia a nonvoter;} \end{cases}$$

$$X_1 = \begin{cases} 1 \text{ if respondent thinks he would be much better off if his favorite cnadidate won than he would be if one of the other candidates won;} \\ 0 \text{ otherwise;} \end{cases}$$

$$X_2 = \begin{cases} 1 \text{ if respondent perceives the election as close; and} \\ 0 \text{ otherwise.}^{251} \end{cases}$$

They first estimate the parameters of their model to obtain variance estimates which they use to correct the heteroscedasticity problem inherent in models with binary dependent variables.[252]

It is not surprising that Ferejohn and Fiorina obtain results in which the coefficient for X_1X_2 is not significantly different from zero, while the coefficient for X_2 is.[253] Closeness is not the only factor in determining probability of affecting the outcome of the election. The size of the electorate also affects this probability.[254] We argue that it is unlikely that the effect of closeness can be detected empirically when the electorate size is large as in modern U.S. Presidential elections. However, when the electorate size is much smaller, effects of closeness can be detected.

Aldrich[255] reviews the competing claims of Riker and Ordeshook[256] on the one hand and Ferejohn and Fiorina[257] on the other. He first notes that the data across elections (i.e., over time) is more supportive of the expected utility hypothesis than data from cross-sections of individuals in the same election, since the respondents have expectations concerning the same event in one election.[258] Using a test of the sort used by Ferejohn and Fiorina, he found support for the contention that the probability term counts when looking at forms of political participation other than voting: voter registration, campaign contributions, speech attendance, etc.[259]

The strategic voting behavior hypothesized by McKelvey and Ordeshook is tested by Black[260] and Cain[261] in two articles that appeared back to back in the August, 1978 issue of the American Journal of Political Science. Using survey data from the 1968 and 1972 Canadian elections, Black finds that a significant proportion of voters cast ballots for their second choice candidates--12 percent in 1968 and 13 percent in 1972.[262] He also finds that the probabilities[263] are important in determining the candidate for whom a voter will cast his ballot.[264] His results are not consistent with the minimax-regret model, particularly since a minimax-regret voter will never vote for his second choice.[265] Black's work only dealt with the voter's choices and their preferences, not participation, but it does support McKelvey and Ordeshook.

Cain used data from England, both individual survey data and aggregate voting data, to test the expected utility hypothesis[266] against the minimax-regret hypothesis.[267] Cain uses the aggregate data for the first test. His model is in the form

$$Y = B_0 + B_1X_1 + B_2X_2 + \varepsilon$$

where:

Y is the percentage of third-party vote in a constituency,

X_1 is the difference between percentage of votes for the winner and he percentage of votes for the runner-up in the constituency,

X_2 is the percentage of eligible voters abstaining, and

ε is the random disturbance term.[268]

His hypothesis is that the closer the race between the major-party candidates, the smaller the percentage of third-party vote, because those voters will not want to waste their votes.[269] Also, he asserts that electors whose favorite candidate is a third-party candidate will be "squeezed" out, making abstention the best alternative, leading to a negative relationship between X_2 and Y.[270] The signs for both B_2 and B_3 are as hypothesized and the estimates are significantly different from zero, supporting the expected utility theory over the minimax-regret theory of abstention.[271]

Using a simultaneous equation maximum-likelihood logit estimation procedure on the survey data, Cain estimated the following equations:

$$V_t = B_{t1} + B_{t2}(P_1U_1 - P_2U_2) + B_{t3}U_1 + B_{t4}(U_1 - U_2) + B_{t5}D + \varepsilon$$

where:

V_t is the three voting alternatives, abstain vote for first choice, vote for a lower choice (vote strategically);

P_1 and P_2 are the reciprocals of closeness to the winner of his first and second choices, respectively, or to the runner-up if his choice was the winner;

U_1 and U_2 are the utilities of the first and second preferences as measured by thermometer scores;

D is a participation index score, an index of involvement in the election and campaign.[272]

The probability that the individual will vote sincerely, $PR(V_1)$, will vote strategically $PR(V_2)$, or will abstain $PR(V_0)$, is expressed as a nonlinear function of the explanatory variables:

$$PR(V_t) = 1 \div e^{-x_t B}.$$

Cain then estimates:

$$\log(PR(V_1)/PR(V_0)) = X\hat{B}_1 \text{ and}$$

$$\log(PR(V_2)/PR(V0\ )) = XB_2,$$

subject to the condition:

$$PR(V_0) + PR(V_1) + PR(V_2) = 1.$$ [273]

Cain seems to have missed the mark in his empirical intepretation of the expected utility term, $P_1U_1 - P_2U_2$, as a predictor of strategic or sincere voting. It should have been $[P_1(U_1 - U_w), P_2(U_2 - U_w)]$, where U_w is the thermometer score of the most viable alternative to candidates 1 or 2. This would have given Cain the expected utility paired comprisons that McKelvey and Ordeshook have demonstrated should be used in the voter's calculate for multicandidate elections.[274]

He finds that his expected-utility term is positively related to the probability of voting sincerely and negatively related to the probability of voting strategically.[275] Only with strategic voting is it a significant factor.[276]

Cain's double-barrelled testing does seem to be more supportive of the expected utility hypothesis than the minimax-regret hypothesis, yet the results of his test seem to be composed of a less-than-efficient mixture of basic components. The simple observation of strategic voting which is not consistent with minimax-regret decision making, should be enough to discredit the minimax-regret hypothesis once and for all.

Finally, in a recent article in Public Choice, Thompson reviews some of the empirical work on rational abstention, with respect to the debate over the decision rule employed by electors, expected utility or minimax regret.[277] He finds that minimax-regret decision-making is not consistent with our observations of people voting for their second choice so as not to "waste" their votes.[278] Minimax regret is also inconsistent with empirical findings that voters participate more in close elections than in landslides.[279] A third prediction of the minimax-regret hypothesis is that turnout in elections should increase with the mean gain resulting from the election, or with "B", yet the evidence from the few studies on this is rather ambiguous.[280] Two things are clear from the literature: first, that closeness counts, and second, that additional articles about minimax regret would only waste journal space.

FOOTNOTES

[1]Lewis Carroll, Alice in Wonderland, edited by Donald J. Gray, New York: W.W. Norton and Company, Inc., 1971, p. 94.

[2]Anthony Downs, An Economic Theory of Democracy, Harper and Row, 1957, pp. 51-53.

[3]Ibid., pp. 36-50, pp. 238-276.

[4]Ibid., pp. 38-45.

[5]Ibid., p. 39.

[6]Ibid., pp. 41-45.

[7]Ibid., pp. 244-245.

[8]Loc. cit.

[9]Loc. cit.

[10]Loc. cit.

[11]Ibid., pp. 260-276.

[12]Ibid., pp. 261-265.

[13]Ibid., p. 265.

[14]Ibid., pp. 267-271.

[15]Ibid., p. 268.

[16]Ibid., pp. 261-262.

17Ibid., p. 265.

[18]Ibid., pp. 260-272.

[19]Ibid., pp. 271-272.

[20]Ibid., pp. 6-11.

[21]Loc. cit.

[22]Gordon Tullock, Toward a Mathematics of Politics, The University of Michigan Press, 1967, p. 114.

[23]William H. Riker and Peter C. Ordeshook, "A Theory of the Calculus of Voting," American Political Science Review, Vol. 62, March, 1968, pp. 27-28.

[24]Op. cit., Tullock, pp. 100-114.

[25]Ibid., pp. 109-114.

[26]Loc. cit.

[27]Tollison and Willet, "Some Simple Economics of Voting and Not Voting," Public Choice, 16, Fall, 1973, pp. 59-71.

[28]Op. cit., Tullock, pp. 109-110.

[29]Loc. cit.

[31]For an example see Riker and Ordeshook, p. 25.

[32]Op. cit., Tullock, p. 114.

[33]However, this is only speculation on our part.

[34]Op. cit., Tullock, p. 114.

[35]Loc. cit.

[36]Op. cit., Riker and Ordeshook, pp. 25-42.

[37]Ibid., p. 25.

[38]Ibid., pp. 27-28.

[39]Op. cit., Tullock, p. 114; Op. cit., pp. 267-271.

[40]Op. cit., p. 28.

[41] For some evidence on these and other reasons that people give for voting or abstaining, there are two survey studies (Harold Foote Gosnell, Getting Out the Vote, Chicago: University of Chicago Press, 1972; and Arthur Twining Hadley, The Empty Polling Booth, Englewood Cliffs, New Jersey: Prentice-Hall, 1978).

[42]Op. cit., Riker and Ordeshook, pp. 27-28

[43]Loc. cit.

[44]Ibid., p. 27.

[45]Op. cit., Tullock, p. 114.

[46]Op. cit., Riker and Ordeshook, pp. 27-28.

[47]Ibid., p. 28.

[48]Loc. cit.

[49]Ibid., pp. 28-34.

[50]Loc. cit.

[51]Loc. cit.

[52]Richard D. McKelvey and Peter C. Ordeshook, "A General Theory of the Calculus of Voting," Mathematical Applications in Political Science, Vol. IV, James F. Herndon and Joseph L. Bernd (eds.), Charlottesville: University of Virginia Press, 1973, pp. 32-78.

[53]Loc. cit.

[54]Ibid., p. 43

[55]Ibid., pp. 32-78.

[56]There were 170 2-member districts, 12 3-member districts, and 1 4-member district in England, Scotland, and Wales in 1868.

[57]The number of votes each elector is allowed to cast is usually less than or equal to the number of winners.

[58]Op. Cit. Tollison and Willett, pp. 59-71.

[59]Martin J. Zechman; "Dynamic Models of the Voter's Decision Calculus: Incorporating Retrospective Considerations into Rational-choice Models of Individual Voting Behavior," Public Choice, 34, 1979, pp. 297-315.

[60]Op. cit., Downs, pp. 38-45.

[61]Nathaniel Beck, "A Note on the Probability of a Tied Election," Public Choice, 23, Fall, 1978, pp. 75-79.

[62]Howard Margolis, "Probability of a Tied Election, Public Choice, Vol. 31, Fall, 1977, pp. 135-738.

[63]I.J. Good and Lawrence S. Mayer, "Estimating the Efficacy of a Vote," Journal of Behavioral Science, 20, 1975, pp. 25-33.

[64]John O. Ledyard, "The Paradox of Voting and Candidate Competition: A General Equilibrium Analysis," California Institute of Technology Social Science Working Paper 224, 1978.

[65]Thomas R. Palfrey and Howard Rosenthal, "Team Games and Voter Participation," Carnegie-Mellon University Graduate School of Industrial Administration Working Paper 25-81-82, 1982.

[66]R. Morris Coats, "Closeness in Elections," Unpublished manuscript, 1982.

[67]Gary Chamberlain and Michael Rothchild, "A Note on the Probability of Lasting a Decisive Vote," Journal of Economic Theory, 25, 1981, pp. 152-162.

[68]Bruno S. Frey, "Why do High Income People Participate more in Politics?", Public Choice, 11 (1971), pp, 101-05.

[69]John Fraser, "Political Participation and Income Level: An Exchange," Public Choice, Vol. 13, 1972, pp. 115-118.

[70]Keith P. Russell, "Political Participation and Income Level: An Exchange," Public Choice, Vol. 13, 1972, pp. 113-114.

[71]Richard G. Niemi, "Costs of Voting and Non-voting," Public Choice, 27, Fall, 1976, pp. 115-119.

[72]Op. cit., Downs, pp. 220-237.

[73]Op. cit., Tullock, pp. 100-114.

[74]Op. cit., Tollison and Willett, pp. 59-71.

[75]Op. cit., Zechman, pp. 297-315.

[76]Op. cit., Tollison and Willett, pp. 62-65.

[77.1]Op. cit., Downs, pp. 240-247.

[78.2]Op. cit., pp. 109-114.

[79]Op. cit., Tollison and Willett, pp. 62-65.

[80]Ibid., pp. 66-67

[81]Otherwise, the predictions of the theory would be contrary to fact.

[82]Op. cit., Zechman, pp. 197-215.

[83]Op. cit., Zechman, pp. 300-312; Op. cit., Downs, pp. 36-50.

[84]Op. cit., Zechman, pp. 304-310.

[85]Ibid., p. 312.

[83]Op. cit., Zechman, pp. 300-312; Op. cit., Downs, pp. 36-50.

[84]Op. cit., Zechman, pp. 304-310.

[85]Ibid., p. 312.

[86]Peter H. Aranson, Melvin J. Hinich, and Peter C. Ordeshook, "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," American Political Science Review, 1, March, 1978, pp. 135-152.

[87]Op. cit., Beck, pp. 75-77.

[88]Ibid., p. 76.

[89]Op. cit., Good and Mayer, pp. 27-30.

[90]Ibid., pp. 28-30.

[91]Loc. cit.

[92]Ibid., p. 30.

[93]Op. cit., Margolis, pp. 135-136.

[94]Ibid., p q 1

[95]Op. cit., Chamberlain and Rothschild, pp. 152-162.

[96]Op. cit., Good and Mayer, p. 30.

[97]Loc. cit.

[98]Loc. cit.

[99]Op. cit., Coats, pp. 6-8.

[100]Op. cit., Palfrey and Rosenthal, pp. 23-25.

[101]Op. cit., Ledyard, pp. 2-3.

[102]Op. cit., Palfrey and Rosenthal, pp. 3-4.

[103]See Frey, Russell and Fraser in "Political Participation and Incomle Level: An Exchange," cited above. Also see Neimi, "Costs of Voting and Non-voting," cited above.

[104]Op. cit., Frey, 1971, pp. 102-103.

[105]Ibid., pp. 102-104.

[106]Ibid., p. 103.

[107]Op. cit., Fraser, pp. 113-114.

[108]Op. cit., Russell, pp. 115-118.

[109]Op. cit., Niemi, pp. 115-119.

[110]Op. cit., Frey, 1971, passim.

[111]Many job offers are of an all-or-nothing nature: if you want the job, you work the hours the boss tells you, no more, no less, or you don't get (or keep) the job. Some jobs may also forbid moonlighting, and the available jobs for moonlighting also involve "lumpy: hours.

[112]Op. cit., Niemi, pp. 115-119.

[113]Ibid., p. 116

[114]Ibid., pp. 116-117.

[115]Loc. cit.

[116]Ibid., pp. 117-118.

[117]Loc. cit.

[118]Our answer is given in terms of dollars, not utils.

[119]John A. Ferejohn and Morris P. Fiorina, "The Paradox of Not Voting: A Decision Theoretic Analysis," American Political Science Review, 66, 1974, pp. 525-536.

[120]See Luce and Raiffa, Games and Decisions, John Wiley & Sons, Inc., 1957, pp. 280-282.

[121]Gerald S. Strom, "On the Apparent Paradox of Participation: A New Proposal," American Political Science Review, 69, 1975, pp. 908-913.

[122]Melvin J. Hinich, "Voting as an Act of Contribution," Public Choice, 35, 1980, pp. 95-100.

[123]Op. cit., Ferejohn and Fiorina, pp. 527-528.

[124]Op. cit., Luce and Raiffa, pp. 278-282.

[125]The regret of a particular action for a particular state of nature being the difference between the highest utility payoff the decision maker could have received if he had known, prior to his decision, what the true state would be.

[126]Op. cit., Luce and Raiffa, p. 280.

[127]Op. cit., Ferejohn and Fiorina, p. 528.

[128]Ibid., p. 534.

[129]Loc. cit.

[130]Ibid., p. 535.

[131]Loc. cit.

[132]Loc. cit.

[133]Lawrence S. Mayer and I.J. Good, "Is Minimax Regret Applicable to Voting Decisions, American Political Science Review, 69, September, 1975, pp. 916-917.

[134]Gordon Tullock, "The Paradox of Not Voting for Oneself," American Political Science Review, 69, September, 1975, p. 919.

[135]Nathaniel Beck, "The Paradox of Minimax Regret," American Political Science Review, 69, September, 1975, p. 918.

[136]Op. cit., Strom, pp. 910-912.

[137]Ibid., p. 911.

[138]Ibid., p. 912.

[139]Ibid., p. 913.

[140]Luce and Raiffa, pp. 275-326 for a discussion of these and other models of decision-making under uncertainty.

[141]Op. cit., Hinich, pp. 95-100.

[142]Several individuals have described their voting behavior to us as trying to vote for the loser, either to help the underdog or to reduce the winner's feeling that he has a mandate.

[143]Stephen J. Brams and Peter C. Fishburn, "Approval Voting," American Political Science Review, Vol. 72, No. 3, September, 1978, pp. 831-847.

[144]Ibid., p. 831.

[145]Ibid., p. 832.

[146]Ibid., pp. 832-837.

[147]Ibid., p. 834.

[148]Loc. cit.

[149]Ibid., p. 833.

[150]Loc. cit.

[151]Ibid., p. 835.

[152]Ibid., p. 837.

[153]Loc. cit.

[154]Loc. cit.

[155]Loc. cit., p. 834.

[156]Loc. cit.

[157]Ibid., pp. 836-837.

[158]Ibid., pp. 837-840.

[159]Ibid., p. 838.

[160]Ibid., p. 832.

[161]Although these are general models, some focus on testing a few key propositions.

[162]Op. cit., Riker and Ordeshook, pp. 34-38.

[163]Richard A. Brody and Benjamin I. Page, "Indifference, Alienation and Rational Decisions," Public Choice, 15, Summer, 1973, pp. 1-17.

[164]Yaram Brazel and Eugene Silverberg, "Is the Act of Voting Rational?" Public Choice, 16, Fall, 1973, pp. 51-58.

[165]Robert Tollison, Mark Crain, and Paul Pautler, "Information and Voting: An Empirical Note," 24, Winter, 1975, pp. 43-59.

[166]Jonathan Silberman and Garey Durden, "The Rational Behavior Theory of Voter Participation: The Evidence from Congresional Elections," Public Choice, 23, Fall, 1975, pp. 101-108.

[167]Russell F. Settle and Buron A. Abrams, "The Determinants of Voter Participation: A More General Model," Public Choice, 27, Fall, 1976, pp. 81-89.

[168]Op. cit., Riker and Ordeshook, pp. 34-38.

[169]Ibid., pp. 34-36.

[170]Ibid., p. 35.

[171]Ibid., pp. 35-36.

[172]Ibid. pp. 36-37. It has been established that this scale or index has a high degree of internal consistency. A. Campbell, G. Gurin, and W.E. Miller, The Voter Decides, (Westport, Conn., 1954), pp. 154-199. Riker and Ordeshook dichotomized the P and B terms into high and low groups. They "tri-chotomized" the D term into high, medium, and low groups.

[173]Op. cit., Riker and Ordeshook, pp. 37-38.

[174]Loc. cit.

[175]Ibid., p. 137.

[176]Ibid., pp. 137-138.

[177]Ibid., p. 137.

[178]Ibid., pp. 137-138.

[179]Ibid., p. 138.

[180]Op. cit., Barzel and Silberberg, pp. 51-58.

[181]Ibid., p. 54.

[182]Ibid., pp. 54-55.

[183]Ibid., p. 54.

[184]Loc. cit.

[185]Ibid., pp. 54-55.

[186]Loc. cit.

[187]Loc. cit.

[188]Loc. cit.

[189]Ibid., p. 56.

[190]Loc. cit.

[191]Loc. cit.

[192]Loc. cit.

[193]Op. cit., Downs, pp. 238-276; Op. cit., Tullock, pp. 100-114.

[194]Op. cit., Tollison and Willett, pp. 54-71.

[195]Op. cit., Tollison, Crain, Pautler, pp. 43-49.

[196]Ibid., p. 46.

[197]Ibid. pp. 46-48.

[198]Loc. cit.

[199]Ibid., p. 47.

[200]Loc. cit.

[201]Had they employed a Leamer analysis of sensitivity to alternative specifications, their results would have been more credible. See Edward E. Leamer's Specification Searches, John Wiley, 1978, pp. 121-201. With all fairness, however, Tollison, Crain and Pautler's work preceded Leamer by three years.

[202]Ibid., p. 45.

[203]See George G. Judge, et al., The Theory and Practice of Econometrics, New York: John Wiley and Sons, 1980, pp. 407-413.

[204]Op. cit., Silberman and Durden, pp. 102-104.

[205]Loc. cit.

[206]Loc. cit.

[207]Loc. cit.

[208]Loc. cit.

[209]Loc. cit.

[210]Loc. cit.

[211]Loc. cit.

[212]Loc. cit.

[213]Ibid., pp. 102-107.

[214]Loc. cit.

[215]Loc. cit.

[216]Loc. cit.

[217]Loc. cit.

[218]Loc. cit.

[219]Loc. cit.

[220]Loc. cit.

[221]Op. cit., Settle and Abrams, pp. 81-89.

[222]Op. cit., Riker and Ordeshook, pp. 34-38.

[223]Op. cit., Tollison, Crain, and Pautler, pp. 43-49.

[224]Op. cit., Silberman and Durden, pp. 101-108.

[225]Op. cit., Barzel and Silberberg, pp. 51-58.

[226]Op. cit., Settle and Abrams, pp. 82-84

[227]Loc. cit.

[228]Loc. cit.

[229]Loc. cit.

[230]Ibid., pp. 84-85.

[231]Ibid., p. 85.

[232]Ibid., p. 87.

[233] See, for example, Phillip E. Converse, "The Concept of the Normal Vote" in Campbell, et. al., Elections and the Political Order, John New York: Wiley and Sons, 1966, p. 24. Also see the discussion of the Burnham-Converse debate on pages 1-3 in the previous chapter and the articles cited there.

[234]Op. cit., Brody and Page, pp. 2-3.

[235]Ibid., pp. 4-6.

[236]Ibid., pp. 5-6.

[237]Ibid., p. 11.

[238]Op. cit., McKelvey and Ordeshook, pp. 42-48.

[239]Leslie Seidle and David Miller, "Turnout, Rational Abstention and Campaign Effort," Public Choice, 27, Fall, 1976, pp. 121-126.

[240]D.T. Denver and H.T.G. Hands, "Marginality and Turnout in British General Elections," British Journal of Political Science, 4, January, 1974, pp. 17-35.

[241]Op. cit., Seidle and Miller, p. 122.

[242]Op. cit., Denver and Hands, p. 33.

[243]Op. cit., Seidle and Miller, pp. 122-123.

[244]Ferejohn and Fiorina, "Closeness Counts in Dancing and Horseshoes," American Political Science Review, 68, 1975, pp. 920-25.

[245]John J. Aldrich, "Some Problems in Testing Two Rational Models of Participation," American Journal of Political Science, 20, November, 1976, pp. 713-733.

[246]Bruce E. Cain, ""Strategic Voting in Britain," American Journal of Political Science, 22, August, 1978, pp. 639-655.

[247]Jerome Black, "The Multicandidate Calculus of Voting: Application to Canadian Federal Elections," American Journal of Political Science, 22, August, 1978, pp. 609-638.

[248]Op. cit., Ferejohn and Fiorina, 1974, pp. 534.

[249]Op. cit., Ferejohn and Fiorina, 1975, pp. 922-924.

[250]ibid., p. 922.

[251]Ibid., p. 923.

[252]Loc. cit.

[253]Ibid., p. 925.

[254]See Chapter 3 in the present work.

[255]Op. cit., Aldrich, pp. 920-25.

[256]Op. cit., Riker and Ordeshook, pp. 25-42.

[257]Ferejohn and Fiorina, "The Paradox of Voting: A Decision Theoretic Analysis," pp. 525-36.

[258]Op. cit., Aldrich, pp. 715-718.

[259]Ibid., pp. 724-726.

[260]Op. cit., Black, pp. 609-638.

[261]Op. cit., Bruce, pp. 639-655.

[262]Op. cit., Black, p. 615.

[263]Objective probabilities are calculated in a manner suggested by McKelvey and Ordeshook - see Black's paper for details.

[264]Op. cit., Black, pp. 624-627.

[265]Op. cit., Ferejohn and Fiorina, p. 534.

[266]Before Cain proceeds with his tests, he develops the expected utility model in a multi-candidate setting, rediscovering (and not citing) McKelvey and Ordeshook's results that had been published six years earlier.

[267]Op. cit., Cain, pp. 644-645.

[268]Loc. cit.

[269]Ibid., p. 645.

[270]Loc. cit.

[271]Loc. cit.

[272]Ibid., pp. 646-651.

[273]Loc. cit.

[274]Op. cit., McKelvey and Ordeshook, pp. 48-52.

[275]Op. cit., Cain, p. 650.

[276]Loc. cit.

[277]Fred Thompson, ""Closeness Counts in Horesehoes and Dancing...and Elections," Public Choice, 38, 1982, pp. 305-316.

[278]Op. cit., Thompson, p. 306.

[279]Thompson argues that although closeness counts, it should account for more of the variation in turnout than it does. He then suggests another alternative to expected utility maximization, prospect theory. In prospect theory "decision weights" replace probability weights and these weights are lower than the probabilities except in the range of low probabilities. Prospect theory seems to be merely a restriction on the form of a utility function.

[280]Op. cit., Thompson, pp. 306-307.

CHAPTER 3

TOWARD A MORE GENERAL THEORY OF THE CALCULUS OF VOTING

I. Introduction

In the previous chapter we surveyed the theoretical literature concerning rational abstention. We found the theory of single-winner, single-vote elections to be well developed from Downs to McKelvey and Ordeshook. Many of the observations in our empirical chapters come from elections in which there are several winners, yet the theory has not been extended to multi-winner, multi-vote elections. This chapter seeks to develop the theory of multi-winner, multi-vote elections.

First we develop the two-candidate, single-winner model of Riker and Ordeshook in Section II. In the third section we extend this model to the multi-candidate case, using McKelvey and Ordeshook's notation. The extension to two-winner, two-vote elections is made in Section IV. In the fifth section the model is generalized and in the sixth section this chapter is summarized.

II. The Individual Choice Model

In the usual presentation of the calculus of voting, the decision to vote or not to vote is considered as a problem of maximizing the expected payoff from the opportunity to vote. If the expected payoff from voting were less than zero, the individual would be better off by not voting at all and would decide to abstain. If the expected payoff were positive, the individual would vote because doing so would increase his utility.

The expected payoff to voting for individual "i," measured in utility, is V_i, given by

3.1 $V_i = P_i B_i + D_i - C_i$;

where:

P_i is individual "i's" subjective probability of affecting the outcome of the election;

B_i is the difference to i in utility between the two possible election outcomes (i.e., between two different winners);

D_i is the utility gain to i from voting that is not associated with i's effect on the outcome; and

C_i is the utility cost to voter i of voting.

Voter i will vote if $V_i > 0$; he will not vote if $V_i < 0$; and he will be indifferent about voting if $V_i = 0$.[1]

A. Factors affecting the variables in V_i

1. The Subjective Probability of Affecting the Outcome (P_i)

If voter i were to have no information about the preferences of fellow voters, he would not have expected one outcome to be any more likely than another outcome, that is, he may assume a uniform distribution of probabilities over the possible outcomes.[2] The probability, P_i, then becomes $1/V$, where V is the number of voters. Voter i may approximate P_i by $1/N$, where N is the number of registered voters, because the voter has no information on the number of non-voting registered voters.

If voter i were to have information about the preferences of other voters from polls or canvasses of the electorate, he would use this information to make a more accurate forecast of the outcome. The closer the polls reveal the two candidates to be, the higher will be the probability of affecting the outcome. Since he has only incomplete information, he will consider the expected outcome (projected by the polls) as most likely and deviations from the expectations of the outcome as less likely as the deviations become larger. Mathematically, P_i can be expressed as

3.2 $P_i = f_i[E(v_1) - E(v_2)]$

where $E(v_1)$ is the proportion of votes projected to be cast for candidate 1 and $E(v_2)$ is the proportion of votes projected to be cast for candidate 2. The function f_i is monotonically decreasing in the absolute value of the difference between the proportion of votes projected for the two candidates.

If the projections have proven unreliable in the past, less weight will be given to the polls, and the probabilities of different outcomes will be closer to a uniform distribution than would be the case if the information were more reliable. So f_i is dependent on the number of registered voters (N) and on how well the polls predict election tallies (G_i). The probability will then be

3.3 $P_i = (1 - G_i)/N + G_i[f_i(E(V_1) - E(V_2)]$,

where G_i is a weighting factor between 0 and 1.[3]

The reciprocal of the number of registered voters times $(1 - G_i)$ becomes an intercept term in the function f_i, and reliability of information becomes the coefficient on the expected vote difference in the probability function.

2. The difference in utility between the candidates (B_i)

The term, B_i, can be viewed as the number difference in i's utility level between having his preferred candidate winning and the other candidate winning. For each i there may be a different B_i. Voters are more likely to view the candidates as being farther apart on the issues if the candidates hold different party allegiances, therefore, B_i will be larger if the candidates are associated with different political parties.

Under a parliamentary system, the party with the majority becomes the party in control of the government. The probability of one's representative making a difference as to which party controls the legislature multiplied by the benefits of having your party instead of the opposition in power is the expected benefit of your candidate changing the parliamentary majority which we will call $p_i b_i$. So B_i is a function of several variables:

3.4 $B_i = (U_1 - U_2) + p_i b_i$,

where U_1 and U_2 are the utilities associated with each candidate's winning of the election, respectively. Notice that if the candidates belong to the same parties, $U_1 - U_2$ will be smaller than if both belong to the same party. Also, $p_i b_i$ will be zero if both candidates belong to the same party.

3. The benefits and costs of voting not associated with the voters effect on the outcome (D_i and C_i)

The term, C_i, represents the sum of the costs of registration, information gathering, transportation to the polls, income and/or leisure foregone, offers of payments for abstaining[5] and any other costs of voting. The term D_i respresents the sum of payments for voting, subsidies of voting costs, benefits of voting because one has a sense of civic duty or because there are social pressures to vote, and any other benefits from voting not associated with changing the outcome.[6]

If an elector is to benefit more from voting than abstaining, the inequality $D_i + P_iB_i > C_i$ must hold. Alternatively, P_iB_i must exceed $C_i - D_i$ for a person to to benefit from voting.

III. Multi-candidate Elections[7]

To examine the calculus of voting more carefully and for extension of the theory to include multi-candidate, multi-winner, multi-vote races, we introduce some notation. The subscripts in the following notation refer to winners or candidates involved in a tie to win, while the superscripts refer to the voting behavior of the citizen under consideration: j if voting for candiate j, and α if abstaining. We have:

m = the number of candidates running in an election,

S = the number of winners in an election,

e = the number of citizens fully qualified (registered) to vote in the election district,

q_k^j = the probability that candidate k wins given that citizen i votes for candidate j,

q_k^α = the probability that candidate k wins given that citizen i abstains,

q_{jk}^j = the probability that candidates j and k tie, given that voter i votes for candidate j,

q_{jk}^α = the probability that candidates j and k tie, given that voter i abstains,

U_j = the utility citizen i associates with candidate j winning the election,[8]

U_{jk} = the utility citizen i associated with candidates j and k tieing to win, and

γ = a set of integers $1 \leq j \leq m$, which identifies a class of voting outcomes, in which candidate j wins in cases where γ contains only; candidates j and k tie to win in cases where γ contains only j and k, etc. The possible sets γ form a partition of the set of all possible voting outcomes.[9]

Since the registered voter can vote for one of the m candidates or abstain, he has m + 1 alternatives to choose among. The attractiveness of a given alternative is determined by the effect that that alternative has upon the outcome and by the utility of the outcome.[11]

A. Assumptions

We will now make several assumptions to simplify our discussion. First, we assume that the citizen is an "outcome taker" that is, the citizen's choice of voting or abstaining does not affect the actions of other citizens.[12] Second, as an approximation, we assume that the probability of a tie among three or more candidates is zero, because the probability of a three-way tie will be roughly the square of the probability of a two-way tie, and the probability of a two-way tie is very small.[13] Third, we assume that $U_1 > U_2 > U_3 > U_4$, etc., without loss of generality, that is, we define candidates in decreasing order of the utility citizen i associates with their victory.

B. The single-winner, two-candidate election

Before we examine multi-candidate elections, let us first examine elections with two candidates and one winner. There are three possibilites for α,

$$\gamma = \begin{cases} \{1\}: & \text{candidate 1 wins,} \\ \{2\}: & \text{candidate 2 wins, and} \\ \{1,2\}: & \text{candidates 1 and 2 tie.} \end{cases}$$

With the single-winner, two-candidate election, we have:

3.5 $q_a^{\alpha} = q_{12}^{1} + q_2^{1}$, and

3.6 $q_1^{1} = q_{12}^{\alpha} + q_1^{\alpha}$.

Equation 3.5 means that the probability that candidate 2 will win, given that voter i abstains, is equal to the probability that candidate 2 will win, even if i votes for candidate 1, plus the probability that candidates 1 and 2 tie if i votes for candidate 1. Equation 3.6 means that the probability that candidate 1 will win, given that voter i votes for 1, is equal to the probability that candidate 1 will win plus the probability that candidates 1 and 2 ties if voter i abstains instead of voting for candidate 1.[14]

To focus on the expected benefit term in equation 3.1, (P_iB_i), we will, for now, disregard the C_i and D_i terms.[15] Defining EB^{α} as the expected benefit from abstaining to the citizen, P_iB_i becomes

3.7 $P_iB_i = EB^1 - EB^{\alpha} = q_{12}^{1}(U_{12} - U_2) + q_{12}^{\alpha}(U_1 - U_{12})$.[16]

The expected utility of voting instead of abstaining is the sum of $q_{12}^{1}(U_{12} - U_2)$ and $q_{12}^{\alpha}(U_1 - U_{12})$.[17] The first term, $q_{12}^{1}(U_{12} - U_2)$, is the difference between the utility of a tie and the utility of a loss times the probability of voter i creating a tie.[18] The second term, $q_{12}^{\alpha}(U_1 - U_{12})$, is the difference between the utility of winning and the utility of tying times the probability of a tie if voter i abstains, which is equal to the probability of breaking a tie if voter i does vote.[19]

With the additional assumption that, if a tie occurs, each candidate has a 50-50 chance of winning (the tie is broken by the toss of a fair coin), U_{12} becomes

3.8 $U_{12} = \dfrac{U_1 + U_2}{2}$.[20]

So, 3.7 can be simplified further by substituting 3.8 into 3.7.[21] Then $EB^1 - EB^\alpha$ becomes

3.9 $EB^1 - EB^\alpha = (q^1_{12} + q^\alpha_{12})(\dfrac{U_1 - U_2}{2})$.[22]

Since probabilities are always greater than zero and U_1 has been assumed to be greater than U_2, then $EB^1 - EB^\alpha \geq 0$. Voter i will vote if and only if $EB^1 - EB^\alpha \geq C_i - D_i$. Since q^α_{12} and q^1_{12} are adjacent probabilities, they are approximately equal, so we will assume $q^\alpha_{12} = q^1_{12} = q_{12}$.[25] Therefore, for m = 2,

3.10 $EB^1 - EB^\alpha = q_{12}(U_1 - U_2) = P_iB_i$.[26]

This is the expected utility term in the original formulation of the vote/abstain decision.[27]

C. The single-winner, three-candidate election

With m = 3 and s = 1, γ has seven possibilities:

γ = <

- {1}: candidate 1 wins,
- {2}: candidate 2 wins,
- {3}: candidate 3 wins,
- {1,2}: candidates 1 and 2 tie for first,
- {1,3}: candidates 1 and 3 tie for first,
- {2,3}: candidates 2 and 3 tie for first, and
- {1,2,3}: all three candidates tie.[28]

Since it is obvious that a citizen will never vote for his least favorite candidate, there are only three rational strategies for the citizen:

Strategy 1: vote for candidate 1,

Strategy 2: vote for candidate 2, and

Strategy 3: abstain.

We are focusing on the expected-benefit term, not whether or not $EB^j - EB^\alpha \geq C_i - D_i$, for the choice between the first two strategies is independent of the size of $(C_i - D_i)$. Therefore, strategy 1 and strategy 2 can be considered separately.

What are the conditions under which strategy 1 is superior to strategy 2? To discover this, we find the difference $EB^1 - EB^2$:

$$3.11 \quad EB^1 - EB^2 = (EB^1 - EB^\alpha) - (EB^2 - EB^\alpha)$$
$$= q_{12}(U_1 - U_2) + q_{13}(U_1 - U_3) - q_{12}(U_2 - U_1) - q_{23}(U_2 - U_3)$$
$$= 2q_{12}(U_1 - U_2) + q_{13}(U_1 - U_3) - q_{23}(U_2 - U_3).$$

Therefore, strategy 1 is to be followed if

$$3.12 \quad 2q_{12}(U_1 - U_2) + q_{13}(U_1 - U_3) > q_{23}(U_2 - U_3),$$

and strategy 2 is to be followed if the inequality is reversed.

The expected utility portion of equation 3.1, $EB^k - EB^\alpha$ in the multi-candidate case, is computed by taking the utility difference between candidate k and each of the other candidates and multiplying each of these differences by a probability term. This probability term is the sum of all of the probabilities of first place ties involving these two candidates.[34] This sum can be interpreted as a measure of competitiveness between the two candidates. Then $EB^k - EB^\alpha$ becomes

3.13 $EB^k - EB^\alpha = \sum_{j \neq k} q^\alpha_{ik} (U_k - U_j)$.

McKelvey and Ordeshook go on to show that the citizen will vote for the candidate who maximizes $EB^k - EB^\alpha$. If no candidate k exists such that $EB^k - EB^\alpha > C_i - D_i$, the citizen abstains.[37]

IV. Multi-winner and Multi-vote Elections

A. Introduction

Some of the districts in the United Kingdom during the nineteenth century returned more than one member from the same race.[40] In a 2-member district, the top two candidates (as far as votes received) would be returned. In both two- and three-member districts, each voter could vote for two of the candidates after 1868.[42] In the one four member district, each voter could vote for three of the candidates.[43] This type of election has not been analyzed in the calculus of voting literature. Before any empirical work is presented, the theory of the calculus of voting must be extended, and made somewhat more general then McKelvey and Ordeshook's "A General Theory of the Calculus of Voting."

Before proceeding to the notation and the analysis of the multi-winner elections, an analogy to horse racing may prove worthwhile. The election-as-a-horse race analogy has been used in the press to describe elections and the dynamic nature of the competition in elections. The only concern for a standard election is in coming in first. Bets at horse races can be made that a given horse will come in first (win), come in either first or second (place), or come in either first, second, or third (show). In some elections, the only thing that is important is to come in first (win), in other elections with two member districts, the only thing that counts is to come in either first or second (win or place); in still other elections, the only thing that counts is to come in either first, second, or third (win, place or show), and so on. In the analysis that follows, we assume that voters care about the ordering of election outcome only to the extent that their preferred candidates are among the returned candidates.

B. Conditional Probabilities

Implicit in the analysis of the single-seat elections in part C of section III is the basic theorem of probability theory that P_r(A and B) = $P_r(A) \cdot P_r$(B given A). For our purposes "A" stands for the event, "candidates j and k are ahead of all other candidates," and "B" stands for the event "candidates j and k tie." In that case P_r(A and B) is equal to the probability that candidates j and k tie for first.

One difference between the single-seat elections and the multi-seat elections can be seen by noticing that the event "A" is different in the two cases. In the multi-seat case "A" stands for the event "candidates i and j are ahead of m-S of the candidates," where m is the number of candidates and S is the number of seats or positions available.

If we expect candiate j, k, ℓ, and m to come in first, second, third, and last, respectively, and the probabilities that two candidates will tie is knife-edged, then the probability that j and ℓ tie will be almost zero compared with the probability that j and k tie or that k and ℓ tie. If the race will result in one winner, then the only probability likely to be large enough to be important is the probability that j and k will tie for first, which will be dependent upon the expected difference in the vote totals for candiates j and k. If the race will result in two winners, then the only probability likely to be important is the probability that candidates k and ℓ will tie for second, which will be dependent upon the expected difference in the vote totals for candidates k and ℓ.

C. Notation

An extension of the notation is needed to extend anaysis to the multi-winner case. For the set of intergers that denote the winners, a dash between two integers means that the candidates represented by those integers are both winners. A comma between two intergers means that the candidates tied for the last winning position in the outcome ordering. Superscripts above variables denote the candidate(s) for whom voter i votes, while the "α" superscript denotes voter i's abstension. For example, with three candidates competiting for two positions, the possibilities for the set "γ" can be represented as:

γ =
- {1--2}: candidates 1 and 2 win, 3 loses,
- {1--3}: candidates 1 and 3 win, 2 loses,
- {2--3}: candidates 2 and 3 win, 1 loses,
- {1--(2,3)}: candidate 1 wins, 2 and 3 tie for second
- {2--(1,3)}: candidate 2 wins, 1 and 3 tie for second
- {3--(1,2)}: candidate 3 wins, 1 and 2 ie for second
- {1,2,3}: all candidates receive the same number of votes (tie for first).[44]

The expected utility terms can be simplified by another assumption and further notation . The probability that candidates j and k will tie for second will be represented as:

q^{α}_{jk} : if voter i abstains,

q^{j}_{jk} : if he votes for candidate j and no other.

q^{jl}_{jk}: if he votes for both candidate j and l.

To make the analysis more manageable, adjacent probabilities are set equal by assumption (as extremely close approximations), as follows:

$$q^{\alpha}_{jk} = q^{j}_{jk} = q^{k}_{jk} = q_{jk}.$$[45]

To simplify the expected benefit term, $EB^j - EB^{\alpha}$, let Δ_{jk} be the utility difference between candidate j and candidate k, that is,

$$\Delta_{jk} = U_j - U_k.$$

D. The Two-Winner, Two-Vote, Three-Candidate Election

In a two-winner, two-vote, three-candidate election, there are several possibilities facing voter i. He can vote for candidate 1, candidate 2, candidate 3, candidates 1 and 2, candidates 1 and 3, candidates 2 and 3, or abstain. It is clear that possibilities involving voting for the least favorite candidate will be dominated by some other strategy, so the obvious possible strategies that voter i may wish to follow, and the associated $EB^{jk} - EB^{\alpha}$ terms are:

Strategy 1. vote only for candidate 1

$$EB^1 - EB^{\alpha} = q_{12}\Delta_{12} + q_{13}\Delta_{13};$$

Strategy 2. vote only for candidate 2

$$EB^2 - EB^{\alpha} = q_{12}(-\Delta_{12}) + q_{23}\Delta_{23};$$ and

Strategy 3: vote for candidated 1 and 2

$$EB^{12} - EB^{\alpha} = q_{13}\Delta_{13} + q_{23}\Delta_{23}.$$[46]

We can now compare the expected benefits among these three alternatives to find out if any strategy is dominated by another strategy (or combination of strategies). We can use these comparisons to state the conditions which must hold for any remaining strategy to be a superior strategy (have a higher $EB^{jk} - EB^{\alpha}$ than all other strategies). To compare any two strategies we simply find the difference between the expected benefit terms. If this difference has an unambiguous sign, then we can tell which strategy is dominated. If there is an ambiguity in the sign of this difference, we can use this to find the conditions for an undominated strategy to be a superior strategy.

First, we compare the expected benefits of strategy 2 with strategy 3. The difference,

$$3.14 \quad EB^{2} - EB^{12} = (EB^{2} - EB^{\alpha}) - (EB^{12} - EB^{\alpha})$$
$$= q_{23}\Delta_{23} - q_{12}\Delta_{12} - q_{13}\Delta_{13} - q_{23}\Delta_{23}$$
$$= -q_{12}\Delta_{12} - q_{13}\Delta_{13},$$

is unambiguously negative. This means that voting only for one's second-favorite candidate is never as good as voting for one's first- and second-favorite candidates (under this particular election structure).

Next, we compare strategy 1 with strategy 3. The difference,

$$3.15 \quad EB^{1} - EB^{12} = (EB^{1} - EB^{\alpha}) - (EB^{12} - EB^{\alpha})$$
$$= q_{12}\Delta_{12} + q_{13}\Delta_{13} - q_{13}\Delta_{13} - q_{23}\Delta_{23}$$
$$= q_{12}\Delta_{12} - q_{23}\Delta_{23},$$

has an ambiguous sign. We can see that strategy 1 is better than strategy 3 iff $q_{12}\Delta_{12} > q_{23}\Delta_{23}$. If $q_{23}\Delta_{23} > q_{12}\Delta_{12}$, then Strategy 3 should be followed.

E. The Two-Winner, Two-Vote, Four-Candidate Election

If we increase the number of candidates to four in a two-winner, two-vote election, the possibilities for γ, the set of possible outcomes, becomes

$$\gamma = \begin{cases} \{1\text{--}2\} \\ \{1\text{--}3\} \\ \{1\text{--}4\} \\ \{2\text{--}3\} \\ \{2\text{--}4\} \\ \{3\text{--}4\} \\ \{1\text{--}(2,\ 3)\} \\ \{1\text{--}(2,\ 4)\} \\ \{1\text{--}(3,\ 4)\} \\ \{2\text{--}(1,\ 3)\} \\ \{2\text{--}(1,\ 4)\} \\ \{2\text{--}(3,\ 4)\} \\ \{3\text{--}(1,\ 2)\} \\ \{3\text{--}(1,\ 4)\} \\ \{3\text{--}(2,\ 4)\} \\ \{4\text{--}(1,\ 2)\} \\ \{4\text{--}(1,\ 3)\} \\ \{4\text{--}(2,\ 3)\} \\ \{1,\ 2,\ 3\} \\ \{1,\ 2,\ 4\} \\ \{1,\ 3,\ 4\} \\ \{2,\ 3,\ 4\} \\ \{1,\ 2,\ 3,\ 4\}. \end{cases}$$

The voting strategies that the citizen may follow (except for abstention and strategies including candidate 4) are:

Strategy 1: plump for candidate 1
$EB^1 - EB^\alpha = q_{12}\Delta_{12} + q_{13}\Delta_{13} + q_{14}\Delta_{14}$,

Stratagy 2: plump for candidate 2
$EB^2 - EB^\alpha = q_{23}\Delta_{23} + q_{24}\Delta_{24} - q_{12}\Delta_{12}$,

Strategy 3: plump for candidate 3
$EB^3 - EB^\alpha = q_{34}\Delta_{34} - q_{13}\Delta_{13} - q_{23}\Delta_{23}$,

Strategy 4: vote for candidates 1 and 2
$EB^{12} - EB^\alpha = q_{13}\Delta_{13} + q_{23}\Delta_{23} + q_{14}\Delta_{14} + q_{24}\Delta_{24}$,

Strategy 5: vote for candidates 1 and 3
$EB^{13} - EB^{\alpha} = q_{12}\Delta_{12} + q_{14}\Delta_{14} - q_{23}\Delta_{23} + q_{34}\Delta_{34}$, and
Strategy 6: vote for candidates 2 and 3
$EB^{23} - EB^{\alpha} = q_{23}\Delta_{23} + q_{34}\Delta_{34} - q_{12}\Delta_{12} - q_{13}\Delta_{13}$

To reduce the number of voting strategies a citizen might rationally employ, and to find the conditions under which he will employ one of those strategies, we proceed as we did in part D, taking the difference in the expected benefits terms between two strategies. First, we ask, "Is voting for candidate 1 and 2 always a better strategy than plumpbing for candidate 2?" The difference in the expected benefits from the two strategies is:

$$3.16 \quad EB^{2} - EB^{12} = (EB^{2} - EB^{\alpha}) - (EB^{12} - EB^{\alpha})$$
$$= q_{23}\Delta_{23} + q_{24}\Delta_{24} - q_{12}\Delta_{12} - q_{13}\Delta_{13} - q_{23}\Delta_{23} - q_{14}\Delta_{14} - q_{24}\Delta_{24}$$
$$= - q_{12}\Delta_{12} - q_{13}\Delta_{13} - q_{14}\Delta_{14},$$

which is always negative. Plumping for one's third choice is also dominated
by the strategy of voting for candidates 1 and 3, that is, $EB^{3} < EB^{13}$

$$3.17 \quad EB^{3} - EB^{13} = (EB^{3} - EB^{\alpha}) - (EB^{13} - EB^{\alpha})$$
$$= -q_{13}\Delta_{13} - q_{23}\Delta_{23} + q_{34}\Delta_{34} - q_{12}\Delta_{12} - q_{14}\Delta_{14} + q_{23}\Delta_{23} - q_{34}\Delta_{34}$$
$$= - q_{13}\Delta_{13} - q_{14}\Delta_{14}.$$

Strategies 2 and 3 above are dominated by strategies 4 and 5 respectively; only voting strategies 1, 4, 5, and 6 remain. Is strategy 6 dominated by strategy 4?

$$3.18 \quad EB^{23} - EB^{12} = (EB^{23} - EB^{\alpha}) - (EB^{12} - EB^{\alpha})$$
$$= q_{12}(-\Delta_{12}) + q_{13}(-\Delta_{13}) + q_{24}\Delta_{24} + q_{34}\Delta_{34}$$
$$- q_{13}\Delta_{13} - q_{23}\Delta_{23} - q_{24}\Delta_{24}$$
$$= - q_{12}\Delta_{12} - 2q_{13}\Delta_{13} - q_{23}\Delta_{23} - q_{14}\Delta_{14} + q_{34}\Delta_{34}.$$

Voting for candidates 1 and 2 will be superior to voting for candidates 2 and 3 iff

$$3.19 \quad q_{12}\Delta_{12} + 2q_{13}\Delta_{13} + q_{23}\Delta_{23} + q_{14}\Delta_{14} > q_{34}\Delta_{34}.$$

Is strategy 6 dominated by strategy 5?

$$3.20 \quad EB^{23} - EB^{13} = (EB^{23} - EB^{\alpha}) - (EB^{13} - EB^{\alpha})$$
$$= q_{12}(-\Delta_{12}) + q_{13}(-\Delta_{13}) + q_{24}\Delta_{24} + q_{34}\Delta_{34} - q_{12}\Delta_{12}$$
$$- q_{14}\Delta_{14} - q_{23}\Delta_{23} - q_{34}\Delta_{34}$$
$$= - 2q_{12}\Delta_{12} - q_{13}\Delta_{13} - q_{14}\Delta_{14} - q_{23}\Delta_{23} + q_{24}\Delta_{24}.$$

Voting for candidates 1 and 3 will be superior to voting for candidate 2 and 3 iff

$$3.21 \quad 2q_{12}\Delta_{12} + q_{13}\Delta_{13} + q_{14}\Delta_{14} + q_{23}\Delta_{23} > q_{24}\Delta_{24}.$$

Is strategy 6 dominated by strategy 1?

$$3.22 \quad EB^{23} - EB^{1} = q_{34}\Delta_{34} + q_{24}\Delta_{24} - 2q_{12}\Delta_{12} - 2q_{13}\Delta_{13}$$

Strategy 6 is not dominated by any of the other three remaining strategies. We have four possible strategies left:

Strategy 1: plump for candidate 1;

Strategy 4: vote for candidates 1 and 2;

Strategy 5: vote for candidates 1 and 3; and

Strategy 6: vote for candidates 2 and 3.

What are the conditions under which each strategy will be employed? The conditions for strategy 6 to be a superior strategy are:

3.23 $q_{34}g\Delta_{34} > q_{12}\Delta_{12} + 2q_{13}\Delta_{13} + q_{23}g\Delta_{23} + q_{14}\Delta_{14}$;

3.24 $q_{24}\Delta_{24} > 2q_{12}\Delta_{12} + q_{13}\Delta_{13} + q_{23}\Delta_{23} + q_{14}\Delta_{14}$;

3.25 $q_{34}\Delta_{34} + q_{24}\Delta_{24} < 2q_{13}\Delta_{13} + 2q_{12}\Delta_{12}$.

The third condition will be satisfied if the first two are satisfied. The first two conditions together are the necessary and sufficient conditions for strategy 6 to be a superior strategy.

What are the conditions for Strategy 5 to be superior? For it to be superior to strategy 1, $EB^{13} - EB^{1} > 0$:

3.26 $$\begin{aligned} EB^{13} - EB^{1} &= (EB^{13} - EB^{\alpha}) - (EB^{1} - EB^{\alpha}) \\ &= q_{12}\Delta_{12} + q_{14}\Delta_{14} + q_{23}(-\Delta_{23}) + q_{34}\Delta_{34} - q_{12}\Delta_{12} \\ &\qquad + q_{13}\Delta_{13} - q_{14}\Delta_{14} \\ &= q_{23}(-\Delta_{23}) + q_{34}\Delta_{34} - q_{1g}\Delta_{13}. \end{aligned}$$

Strategy 5 is preferred to strategy 1 if and only is $q_{34}\Delta_{34} > q_{23}\Delta_{23} + q_{13}\Delta_{13}$. For strategy 5 to be superior to strategy 4, $EB^{13} - EB^{12}$ must be positive:

3.27 $$\begin{aligned} EB^{13} - EB^{12} &= (EB^{13} - EB^{\alpha}) - (Eb^{12} - EB^{\alpha}) \\ &= q_{12}\Delta_{12} + q_{14}\Delta_{14} + q_{23}(-\Delta_{23}) + q_{34}\Delta_{34} - q_{13}\Delta_{13} \\ &\qquad - q_{23}\Delta_{23} - q_{14}\Delta_{14} - q_{24}\Delta_{24} \\ &= q_{12}\Delta_{12} + q_{34}\Delta_{34} - 2q_{23}\Delta_{23} - q_{13}\Delta_{13} - q_{24}\Delta_{24}. \end{aligned}$$

Strategy 5 is preferred to strategy 4 iff

3.28 $q_{12}\Delta_{12} + q_{34}\Delta_{34} > 2q_{23}\Delta_{23} + q_{13}\Delta_{13} + q_{24}\Delta_{24}$.

Strategy 5 is preferred to strategy 6 iff:

3.29 $q_{24}\Delta_{24} < 2q_{12}\Delta_{12} + q_{13}\Delta_{13} + q_{23}\Delta_{23} + q_{14}\Delta_{14}$

as we can see from inequality 3.21. The conditions for Strategy 5 to be a superior strategy are then:

3.30 $q_{34}\Delta_{34} > q_{23}\Delta_{23} + q_{13}\Delta_{13}$ (to be preferred to 1);

3.31 $q_{12}\Delta_{12} + q_{34}\Delta_{34} > 2q_{23}\Delta_{23} + q_{13}\Delta_{13} + q_{24}Q_{24}$ (to be preferred to 4); and

3.32 $q_{2g}\Delta_{24} < 2q_{12}\Delta_{12} + q_{13}\Delta_{13} + q_{23}\Delta_{23} + q_{1g}\Delta_{14}$ (to be preferred to 6).

Notice that when strategy 6 is preferred to strategy 4, strategy 5 is also preferred to strategy 1.

To the conditions for strategy 4 to be a superior strategy, we must first find the conditions for it to be preferred to strategy 1, that is $EB^{12} - EB^{1} > 0$:

3.33 $EB^{12} - EB^{1} = q_{23}\Delta_{23} + q_{24}\Delta_{24} - q_{12}\Delta_{12}$.

So strategy 4 is preferred to strategy 1 if $q_{23}\Delta_{23} + q_{24}\Delta_{24} > q_{12}\Delta_{12}$. Strategy 4 is a superior strategy if, in addition to this condition, the following conditions are also met:

3.34 $2q_{23}\Delta_{23} + q_{13}\Delta_{13} + q_{24}\Delta_{24} > q_{12}\Delta_{12} + q_{34}\Delta_{34}$ (to be preferred to 5--from inequality 3.28); and

3.35 $q_{12}\Delta_{12} + 2q_{13}\Delta_{13} + q_{23}\Delta_{23} + q_{14}\Delta_{14} + q_{34}\Delta_{34}$ (to be preferred to 6--from inequality 3.19).

F. The 2-seat, 2-vote, n-candidate election

What happens to the analysis when the number of candidates in a two-seat, two-vote election is m? The possibilities for the outcome set, γ, is

γ = <

{1--2}
|
{1--m}
{1--3}
|
{(m-1)--m}
|
{1---2,3}
|
{1--2,m}
|
{m--(m-2), (m-1)}
|
{1--2,3,4}
|
{1--2,3,4}
|
etc.
{1,2,3,...,m}.

The possible expected utilities of altering the outcome by voting can be expressed as $EB^j - EB^\alpha$ for all plumper votes and $EB^{jk} - EB^\alpha$ for all two-candidate votes. The relevant probability terms can be written in the form:

q_{jk} = the probability that voter i breaks (or creates) a tie between candidates j and k.

The relevant utility terms for plumping for candidate j are:

Δ_{jk} = the utility of breaking a tie between candidates j and k for second, or the utility of creating a tie between candidates j and k for second.

The expected benefits of plumping for candidate j (voting only for j) can now be written as:

3.36 $EB^{j} - EB^{\alpha} = \sum_{k \neq j} q_{jk}\Delta_{jk}$

V. Generalizing the Model

A. Introduction

The model and the analysis in the previous section can easily be extended in two directions: to cases where there are more than two winners and to cases where voters can cast more than two votes. In the first class of elections, three or more winners, we will have the same basic analysis with one slight exception, the probability of a tie between candidates j and k are conditional on the probability that S-1 of the candidates will get more votes than j or k. Included in the second class of elections is approval voting.

Before we generalize our results for this second class of elections, we must define one term. Let us first define an allowable strategy set as a set containing candidates, which if a voter marks the names of all candidates in that set on his ballot, election officials will count his ballot.

B. Propositions

Results from the previous sections can now be generalized for all non-weighted, single-ballot elections as the following propositions:[47]

Proposition 1: If a voter is voting for fewer candidates than is allowable, or if the number of allowable votes is one less than the number of candidates, the voter votes for his favorite candidate;

Proposition 2: If $U_b > U_j$ for all candidate j ≠ b, and there exist some candidates, d and c, such that $U_b > U_d > U_g$, then a voting strategy, T, which does not contain candidate d but does contain candidates b and g, can not be dominated by a strategy, S, is the strategy T augmented by the candidate d; and

Proposition 3: If a voter's strategy set containing the largest allowable number of candidates does not contain the voter's least favorite candidate, that strategy can not be dominated by another strategy which contains the same number of elements.

C. Proof of Proposition 1

Let us first introduce the following notation:

C: a strategy set of candidate with n elements;

c: any candidate, $c \in C$;

b: voter's favorite candidate $b \notin C$

B: a strategy set of candidates with n+1 elements which contains all candidates, $c \in C$ (i.e., $C \subset B$) and candidate $b \notin C$, where b is the $(n+1)^{th}$ element of B;

A: a set of candidates containing all candidates in the election except those candidates in B, i.e., $A = \sim B$;

a: any candidate $a \in A$;

n: the number of candidate in C;

m: the number of candidates in the election; and

v: the number of votes a voter is allowed to cast in the race.

The expected benefits of voting for the candidates in B and C instead of abstaining will be denoted by EB^B and EB^C, respectively. To compare the voter's strategy of voting for candidates in C with the strategy of voting for candidates in B, we substract EB^B from EB^C,

$$3.37 \quad EB^C - EB^B = (EB^C - EB^\alpha) - (EB^B - EB^\alpha)$$

$$= \sum_{a\varepsilon A} \sum_{c\varepsilon C} q_{ca}\Delta_{ca}$$

$$+ \sum_{c\varepsilon C} q_{cb}\Delta_{cb} - \sum_{a\varepsilon A} \sum_{c\varepsilon C} q_{ca}\Delta_{ca} - \sum_{a\varepsilon A} q_{ba}\Delta_{ba}$$

$$= \sum_{c\varepsilon C} q_{cb}\Delta_{cb} - \sum_{a\varepsilon A} q_{ba}\Delta_{ba}$$

If candidate b is voter's favorite candidate, i.e., $U_b > U_c$ for all $c \varepsilon C$ and $U_b > U_a$ for all $a \varepsilon A$, and if strategy B is allowable, i.e., $n + 1 \leqslant v$, then strategy B dominates strategy C, because $\Delta_{cb} < 0$ for all c.

D. Proof of Proposition 2

Extending the notation in part C, above, we have:

T: a strategy set containing n elements, including b;

t: any element of T, $t \varepsilon T$;

d: a candidate $d \not\varepsilon T$;

S: a strategy set containing n+1 elements, n elements are the elements T, and the $(n+t)^{th}$ element is candidate d;

R: a strategy set containing m-n elements such that R is the complement of T, i.e., $R = \sim T$;

r: a candidate $r \varepsilon R$; and

n: the number of candidates in T.

The expected benefits of voting for all candidates in T and the expected benefits of voting for all candidate in S will be denoted by EB^T and EB^S, respectively. To compare the voter's strategy of voting for all candidates in T with the strategy of voting for all candidates in S, we subtract EB^S from EB^T.

3.38 $EB^T - EB^S = (EB^T - EB^\alpha) - (EB^S - EB^\alpha)$

$$= \sum_{r\epsilon R} \sum_{t\epsilon T} q_{tr}\Delta_{tr} + \sum_{t\epsilon T} q_{td}\Delta_{td} - \sum_{r\epsilon R} \sum_{t\epsilon T} q_{tr}\Delta_{tr}$$

$$- \sum_{r\epsilon R} q_{dr}\Delta_{dr}$$

$$= \sum_{t\epsilon T} q_{td}\Delta_{td} - \sum_{r\epsilon R} q_{dr}\Delta_{dr}.$$

Since there exists at least one $t \in T$, such that $U_t > U_d$ (where $t = b$), and at least one $t \in T$, such that $U_t < U_d$ (where t is any $t \neq b$), the sign of $EB^T - EB^S$ is ambiguous, so T can not be dominated by S.

E. Proof of the Third Proposition

To prove the third proposition let us define the following sets of candidates and their elements:

G = a set of candidates containing v-1 elements, $g \in G$,

$H = \{h\}$, $h \notin G$

$I = G \cup H$, $i \in I$,

$J = \{j\}$, $j \notin G$

$K = G \cup J$, $k \in K$,

$M = \sim I$, $m \in M$.

$N = \sim K$, $n \in N$,

$L = \sim G = M \cup H = N \cup J$, $\ell \in L$, and

$O = L \cap M \cap N$, $o \in O$.

Further, let us assume that $U_j > U_a > U_g$ for all $g \in G$ and that $U_g > U_o$ for all $g \in G$ and for all $o \in O$.

We shall be comparing the two strategy sets, I and K. Neither set contains the voter's least favorite candidate. Both sets contain the elements in G. Set K also contains candidate j (j could even be the voter's favorite), while set I contains, in addition to the candidates in G, candidate hH. To compare the two strategies, we subtract the expected value of voting for all candidates in K from the expected value of voting for all candidates in I.

$$3.39 \quad EB^I - EB^K = (EB^I - EB^\alpha) - (EB^K - EB^\alpha)$$

$$= \sum_{i\varepsilon I} \sum_{m\varepsilon M} q_{im}\Delta_{im} - \sum_{k\varepsilon K} \sum_{n\varepsilon N} q_{kn}\Delta_{kn}$$

$$= \sum_{g\varepsilon G} \sum_{o\varepsilon O} q_{go}\Delta_{go} + \sum_{g\varepsilon G} q_{gj}\Delta_{gj}$$

$$+ q_{hj}\Delta_{hj} - \sum_{g\varepsilon G} \sum_{o\varepsilon O} q_{go}\Delta_{go}$$

$$- \sum_{o\varepsilon O} q_{jo}\Delta_{jo} - \sum_{g\varepsilon G} q_{gh}\Delta_{gh} - q_{jh}\Delta_{jh}$$

$$= \sum_{o\varepsilon O} q_{ho}\Delta_{ho} + \sum_{g\varepsilon G} q_{gj}\Delta_{gj} - \sum_{O\varepsilon O} q_{jo}\Delta_{jo}$$

$$- \sum_{g\varepsilon G} q_{gh}\Delta_{gh} - 2q_{jh}\Delta_{jh}.$$

Since $U_j > U_g$ for all g, and $U_j > U_h$, $\Delta_{gj} < 0$ and $\Delta_{jh} > 0$. Strategy I will be superior to strategy K if and only if

$$3.40 \quad \sum_{o\varepsilon O} q_{ho}\Delta_{ho} > \sum_{g\varepsilon G} q_{jg}\Delta_{jg} + \sum_{o\varepsilon O} q_{jo}\Delta_{jo} + \sum_{g\varepsilon G} q_{gh}\Delta_{gh} + 2q_{jh}\Delta_{jh},$$

which is possible.

VI. Summary

In this chapter we have extended the basic analysis of McKelvey and Ordeshook on multi-candidate elections to elections with multiple winners and to elections which allow the voters multiple votes (although not multiple votes for the same candidate). In our extension to multiple winners, we found that the primary focus of the election moves with the number of winners from a focus between the candidates expected to be the first and second vote getters in a one-winner race to a focus between the candidates expected to be the second and third in a two-winner race. In our extension to the class of all non-weighted, single-ballot elections, we proved these three propositions:

1. If a voter is voting for fewer candidates than is allowable, or if the number of allowable votes is one less than the number of candidates, the voter votes for his favorite candidate.

2. If $U_b > U_j$ for all candidates $j \neq b$, and $U_b > U_d > U_{gc}$ for at least some candidate d and candidate g, then a voting strategy, T, which does not contain candidate d but does contain candidates b and g can not be dominated by a strategy, S, which contains candidate d and all elements of T.

3. If a voter's strategy set containing the largest allowable number of candidates does not contain the voter's least favorite candidate, that strategy can not be dominated by any other stratery set of the same size.

Our first proposition shows that under approval voting, a voter will mark the name of his favorite candidate on his ballot. Our second proposition states that a voting strategy set, T, containing a voter's favorite candidate and a third or lower favorite (except his least favorite) can, under certain circumstances, be superior to a strategy containing all of the original candidates in T and a candidate that is not as preferred as his favorite but preferred to some of the other candidates in T. Our third proposition states that if two strategy sets of the same size, S and T, strategy T contains all elements of S except for one candidate, strategy S can not dominate strategy T. These propositions are not new results. Our approach to the proofs has the advantage of simplicity, however.

FOOTNOTES

[1]Op. cit., Riker and Ordeshook, p. 25

[2]He may however, use a decision rule other than the equi-probability-outcomes expected utility rule. See Luce and Raifff, pp. 275-294.

[3]Op. cit., Tullock, p. 109. This G_i term is much like Tullock's A term.

[4]Loc. cit.

[5]Op. cit., Cox and Kousser, 1979, passim.

[6]Loc. cit.

[7]This entire section is based on the work of McKelvey and Ordeshook, cited previously.

[8]This assumes that the utility of an election outcome is independent of the majority size, that is, the voter is concerned with the identity of the winner and not the margin of victory.

[9]Op. cit., McKelvey and Ordeshook, passim.

[11]Ibid., p. 35

[12]Ibid., p. 38.

[13]Ibid., pp. 51.

[14]Ibid., pp. 40-41.

[15]The choice among non-abstaining strategies is independent of both C_i and D_i.

[16]Op. cit., McKelvey and Ordeshook, p. 41.

[17]Loc. cit.

[18]Loc. cit.

[19]Loc. cit.

[20]Ibid., p. 42

[21]Loc. cit.

[22]Loc. cit.

[25]Loc. cit.

[26]Loc. cit.

[27]Loc. cit.

[28]Ibid., p. 43.

[33]Ibid. , p. 51

[34]Loc. cit.

[35]Loc. cit.

[36]Loc. cit.

[37]Loc. cit.

[38]Ibid., p. 52.

[39]Loc. cit.

[40]Craig. passim.

[41]Loc. cit.

[42]Ibid, p. xvii.

[43]Loc. cit.

[44]Although this is one of the possibilities, and therefore a possible element of γ, its occurrence is so unlikely that we will assume that ties of more than two candidates will occur with probability equal to zero.

[45]The objective probability functions we are dealing with are multinomial functions which, although discontinuous, have "jumps" that are quite small relative to the size of the probabilities. Our assumption that adjacent probabilities are equal is reasonable and is made by McKelvey and Ordeshook in their work on page 42.

[46]$EB^3 - EB^\alpha$ is negative and always less than $EB^{jk} - EB^\alpha$ for all other j's and k's.

[47]As mentioned in the previous chapter, these propositions can be considered as corrolaries to a theorem by Brahms and Fishburn.

CHAPTER 4

DEVELOPING THE EMPIRICAL MODEL

I. Introduction

After having shown that the expected closeness of the candidates in terms of votes and that the size of the electorate will still affect the probability of altering the election outcome even in multi-winner situations, we can begin the process of building an empirical version of our model of voter participation. In the next section we deal with the problem of heteroscedasticity that is inherent in the binary nature of the vote/abstain decision. The data set chosen for this study and the basic variables used in the empirical model are discussed in the third section of this chapter. The important variables are separated from the insignificant variables and the functional forms of the variables are honed in the fourth section, to give us the model we will use for testing purposes. In the fifth section, the coefficients for the 1880 unweighted model and the 1865, 1868, and 1874 weighted models are interpreted. The final section summarizes the results of this chapter.

II. Quantal Choice and Zellner's Seemingly Unrelated Regression Model

Although some of the variables affecting the individual's choice to abstain or vote are continuous, the observable dependent variable, vote or abstain, is binary. This model can be represented generally as

4.1 $y_i = f(X_i) + \varepsilon_i$

or for OLS as

4.2 $y_i = X_iB + e_i, \; i = 1,\ldots,M$

where y_i is a variable that takes the value of 1 if citizen i votes and 0 if he abstains, X_i is the vector of independent variables, B is the unknown vector of parameters, and e_i is the random disturbance term.

It is well known that OLS can not be directly applied because its assumption of homoscedastic errors is violated. For $E(e_i) = 0$, e_i can take on only two values, $-(X_iB)$ or $1-(X_iB)$, which means that the error terms take on these values with probabilities $1-(X_iB)$ and X_iB, respectively.[1] A further problem is that X_iB, the prediction of y_i, can have values outside the unit interval.[2] Finally, the variance of e_i will be

4.3 $\quad Var(e_i) = (X_iB)(1-X_iB)$

because of the Bernoulli distribution of y_i.[3]

Several other problems occur with the use of OLS to estimate the parameters of a binary choice model. One difficulty is that if values for X_i are outside of the range of the sample, the predicted y_i's can be outside the unit interval, even if in-sample predictions of y_i fall within the unit interval.[4] Still another objection is that any method of estimation that is linear in the y_i's is not generally efficient, since the y_i's are not normally distributed.[5]

Zellner and Lee have suggested an approach to correct for the heteroscedasticity when the values for the X_i vector are identical (i.e., several observations have common values for the entire X_i vector).[6] Then we can aggregate the model in (4.2) into

4.2 $\quad P_j = X_jb + \varepsilon_j.$

If E_j is the number of registered voters in dstrict then $P_j = \sum_{i=1}^{E_j} \frac{y_i}{E_j}$ is the proportion of registered voters in district j who actually voted; for all citizens in district j, $X_{ij} = X_j$ the vector of independent variables for all citizens in district j;

b is the vector of coefficients relating the macro-variables to P_j's and;

ε_j is the vector of errors with mean zero and variance

$$\left(\frac{\overline{P}_j(1 - \overline{P}_j)}{E_j}\right) = X_j b \text{ where } \overline{P}_j = E(P_j).^7$$

In our data set, we have observations not just of different districts at one time, but over several elections (several time periods). To apply the joint-estimation procedure suggestion by Zellner and Lee, we would need estimates of the proportion of the electorate in a district that voted in both elections, which is not available. We can, however, procede with joint estimation in another way, assuming there is serial correlation of errors over time in a single district.

One reason for serial correlation is that our model is likely to be incomplete, i.e., it does not contain all relevant variables. We can see this as

$$4.5 \quad P_j = X_{j1}B_1 + X_{j2}B2 \quad + \varepsilon_j,$$

where ε_j is distributed as in 4.2;

X_{j1} is the vector of variables that are included; B_1 is the vector of coefficients for the included variables; X_{j2} is the vector of variables not included (perhaps because of insufficient data); and B_2 is the vector of coefficients for the excluded variables.

The X_{j2} variables are local characteristics, characteristics peculiar to district j. These characteristics are generally slow to change, so that both X_{j2} and B_2 are assumed to be stable (B_2 represents the voter's reaction to X_{j2}). Thus we assume $E(X_{j2t}B_{2t}) = X_{j2t-1}B_{2t-1}$ where t is the time period index.

We can write 4.4 as

4.6 $P_{jt} = X_{1t}B_{1t} + u_{jt}$

where $u_{jt} = X_{j2t}B_{2t} + \varepsilon_{jt}$. If $X_{j2}B_2$ is persistent over the time, then the expected value of u_{jt} would be

4.7 $E(u_{jt}) = X_{j2t}B_{2t}$,

but X_{j2} is not observed so the expected value of $X_{j2t}B_{2t}$ can be used, which is

4.8 $E(u_{jt}) = X_{j2t-1}B_{2t-1}$.

Since

4.9 $E(X_{j2t-1}B_{2t-1}) = u_{jt-1}$,

4.8 becomes

4.10 $E(u_{jt}) = u_{jt-1}$.

We can use 4.10 to obtain an Aitken estimator for B_1. This is a form of Zellner's SUR which makes use of the autocorrelation of the errors. If an observation has no related observation in the previous (or subsequent) election, a neutral weight the means of the weighting values, can be assigned to that observation.

We will use weighted least squares for another reason that has not been mentioned yet. To determine whether the secret ballot had an effect, we use a Chow test, a test for structural change. It is well known that if the errors are heteroscedastic, the Chow test does not reject the hypothesis that two regression samples were drawn from the same population when that hypothesis should be rejected. As Toyoda[8] points out, the use of GLS to correct for heteroscedasticity improves the power of the Chow test.

III. The Data Set

Among the objectives of this work is to show that the size of the electorate (E) does affect the probability of a voter changing the ordering of the candidates (Pr). This can only be done empirically by showing the effect of electorate size on voter participation, which requires a data set in which the variance of the electorate sizes is large. If the variance of the electorate sizes is small, then the effect of electorate size on turnout may not be discernable. Also, if a large electorate size does wash out the effect of predicted closeness on the probability of affecting the ordering, then the data set must contain some observations with very small electorates if it is going to detect a closeness effect. The data set used by Ferejohn and Fiorina in their test of the effect of closeness on participation contained no small electorates, and electorate size had a small variance.[9]

Another objective of this paper is to test the hypothesis offered by Cox and Kousser concerning the effect of a change in the election institutions (the introduction of the Australian ballot) on turnout. The data set must contain observations that are under secret balloting and other observations under open balloting.

The data set chosen was British Parliamentary election results from 1860 to 1885 found in F.W.S Craig's British Parliamentary Election Results 1832-1885. This source gave information on electorate sizes;

candidates' names, parties, and votes; the occurrence of boundary changes; the number of seats in the district and the number of votes each voter can cast. It also classified each district according to location (e.g., London, Wales, Scotland) and type of district (borough or county). The geographic area of districts in square miles was found in several sources: British Sessional Papers, 1867/68 Vol. XX; Dod's Electoral Facts, 1832-53; and McCalmont's Parliamentary Poll Book of all Elections, 1832-1918. Data on population[10] and property value subject to the income and profits taxes was found in McCalmont's Parliamentary Poll Book.

As Craig points out in the "Introductory Notes" to his work, due to the system of voter registration in England during the period, some electors have their names recorded more than once in the same electoral registrar.[11] Therefore, the electorate figures should be regarded as approximations. Also, in multi-seat districts voters could cast more than one vote, but sometimes would cast "plumpers," voting for only one of the candidates to put that candidate in a better position relative to the rest of the pack.[12] In computing turnout (T) in multi-seat districts, the number of votes cast is divided by the product of electorate size and the number of votes that could be cast by each elector in that district.

Information is not available on the subjective probability of affecting the ordering, benefits of affecting the outcome, costs of voting, benefits of voting not associated with altering the ordering, and so forth. This means that proxies must be used.[13]

Variables related to the probability of changing the ordering are closeness (CL), and electorate size (E). As sugested in Chapter III, the expected closeness of the votes between the last winner and the first loser should have a positive effect on the probability of affecting the election outcome. Our closeness variable, CL, is $\frac{V_1 - V_2}{E_j}$, where V_1 represents the actual votes cast for the last winner and V_2 represents the actual votes cast for the first loser in the ordering. Note that CL varies inversely with the closeness of the candidates so that its effect on turnout should be negative. Although one-hundred percent canvassing was actually done and the results of the canvas were widely known, these canvass results are not now available.[14] Actual vote percentages, given that supporters of the various candidates are equally likely to vote, are good estimates of the expected percentages. Electorate size, E, should also have a negative impact on turnout because of its negative effect on the probability of affecting the outcome, as suggested in Chapters 3.

Another important term in voter's calculus equation is B, the benefits of changing the outcome. Our proxy for B is a dummy variable we call party change (PCH). If the last winner (j) and first loser (k) belong to different political parties, the utility difference, Δ_{jk}, will be larger because their positions on the issues will be more divergent than the positions of two candidates from the same party.

Another major term in the calculus of voting is the cost of voting (C) which includes deciding who to vote for. One major cost of voting in the nineteenth century was the transportation costs to the polls. We tried three proxies for this: a borough/county binary variable

(COUNTY) which has a value of one if the district is a county and zero otherwise; the square root of the area of the district (to approximate average distance between two points in the district which we call AREA; and population density (DENS) because the more densely populated a district is the closer the polling place is likely to be to any voter.

As Bruno Frey suggests, there is an effect of income on participation. The best proxy we found for income was the per capita property value (PROPAV) for each district. Another possible variable is the number of 10£ houses in the district (HOUS).[15]

The last term in the voter's calculus is the benefit from voting not associated with changing the election outcome, D. One type of benefit not associated with affecting the outcome is being paid to vote for a particular candidate. A proxy for this activity is whether or not an election petition (a charge of election fraud) is brought against one of the candidates. This may not be an adequate measure of bribery, since the loser would have no incentive to bring charges if he were also involved in bribery.

Three dummy variables are included to control for "cultural" differences, differences in voter ethics, etc., which would be included in the D term, as well. One is a dummy for Scottish districts (SCOT). The Scottish dummy is included because of the "overwhelmingly liberal character of Scottish boroughs" and counties,[16] and because of the close social structure (greater social pressure to vote, or in other words, a larger D for all Scottish voters, hence higher turnout in Scotland). Another cultural dummy is for Welsh districts (WAL). This variable is included because of the difference in social structure that the Welsh have from both the English and the Scots. There is no expectation on the sign of the coefficient for this variable. The third dummy is for London borough districts (LOND). These districts were usually much larger in electorate size than most other districts (on the order of 10 times larger). The following is a description of the London boroughs, appearing originally in the Law Times in 1868.

> In London, the electors are so numerous that nobody values the franchise. Who cares to be one of fifty thousand? The consequence is, that one half of the constituency never goes to the poll at all. Three-fourths of the entire body are altogether indifferent to politics, and will not trouble themselves to walk across the street to record their votes. It matters not to them whether Mr. A. or Mr. B. is returned; so the election is really left to the candidates who must exert themselves to bring the voters to the poll. It is all an affair of organization and cost. How are the electors to be induced to vote for you? Canvass is out of the question; they are careless of politics, with the exception of a few noisy cliques. The first step, then, and the only one, is to form committees in every district of persons known in that district, who will persuade the voters to go to the polls, not for the sake of the candidate, but to oblige their neighbor, the committeeman. These committees must hold their meetings at public-houses, and eat and drink, and employ assistants, and many of themselves expect payment for their services. But even to oblige a committee no voter will walk to the poll, though it be but five hundred yards distant. He must be taken to the poll in a cab. Consequently, the candidate is compelled to hire for the day all the cabs he can procure. The mere sending round of the address by the penny costs 500 pounds. It will thus be seen how it comes that many of the metropolitan elections cost 12,000 pounds, and few fall below 6000 pounds.[17]

The last variable we use is unclassified in terms of the voter's calculus. If we were to put it anywhere it would be with the "D" variables. This variable is the percentage change in electorate size (number of registered voters) since the last general election, electorate change (ECH). This is to control for the increase in new voters. The effect here is ambiguous, since it affects different voters differently. If a citizen goes to the trouble of registering to vote, he is likely to vote in the election immediately after he registers. If the franchise is greatly extended and these people register, by far most of these will vote, which is likely to increase turnout, though turnout among those who long held the franchise may decrease because of the dilution of the franchise. An increase in ECH may cause turnout to increase it since voting gave the individual greater status in the community. For many of our observations ECH could not be computed because the district was newly created or because of a change in its boundaries. These observations were deleted from our data set.

IV. Choice of the Design Matrix

In applied econometric work, we seem to be between a rock and a hard place. We need the data for hypothesis testing, yet our tests will be misleading if we do not correctly specify the design matrix. However, it is hard to specify the design matrix correctly without using the data to guide us in our search. Use of a data set to specify the model (i.e., to discover the "true" model) will lead to estimates with unknown sampling properties, rendering that data set impotent in hypothesis testing.[18]

One way out of this difficult position is to use one data set for the search procedure and another data set for estimation and hypothesis testing. We use data from the 1880 British Parliamentary elections for the search. Then, the parameters for the 1868 and 1874 elections will be estimated, using the 1880 residuals in the weighting procedure described above to weight the corresponding 1874 observations and the resulting 1874 residuals to weight the corresponding 1868 observations.

Our search procedure has three distinct steps. In the first step, we eliminate independent variables from the model that have no significant impact on our dependent variable using a backward elimination regression technique to find the "right variables." In the next step, we compare various algebraic forms (non-linear) for expressing the quantitative variables that survived the first step of the search. In the last step, we repeat the first step, including all of the previously eliminated insignificant variables (some of these variables may become significant with the stronger algebraic forms of the quantitative variables).

With the backward elimination procedure, we begin with all variables that we think may have an effect on turnout included in a regression equation. With each successive run, the regressor contributing the least to the model (in increasing R^2) is eliminated until all variables are significant at the .10 level.[19] The method is one generated by SAS. The automatic characteristic of the method inhibits manipulation by the researcher, although the procedure is somewhat *ad hoc*.

Kennedy and Bancroft, using a numerical study, report that the backward elimination stepwise regression method is relatively more efficient[20] then the forward addition stepwise regression method, another automatic search procedure. A reason for this may be that the forward selection procedure begins with biased estimators and may incorrectly choose to add the wrong variables. The backward deletion procedure, on the other hand, begins with unbiased (but not least-variance) estimators.

We begin our "search for truth" with a backward stepwise elimination regression procedure using data from the 1880 election. The dependent variable, turnout (T), is regressed on closeness (PRM), electorate size (E), the percentage change in electorate size from the last election (ECH), the party-change dummy (PCH), the county/borough dummy (COUNTY), the London dummy (LOND), the Scottish dummy (SCOT), the Welsh dummy (WAL), the petition dummy (PET), the square root of area (AREA), per capita property value (PROPAV), the population density (DENS), and the number of inhabited houses (HOUS).

The results of the backward stepwise search are given in Table IV.1 of this chapter. The estimated coefficients for AREA, HOUS, PROPAV, DENS, PET, and ECH are not significantly different from zero at the .10 level and so these variables are removed from the model. The first step in our search is complete.

The second step in our search is to try alternative algebraic forms of our quantitative independent variables--the relationships may be curvilinear. We will first try different forms for the base of the closeness term, then for the different forms of the electorate size, and then different forms for the closeness term again.[21]

TABLE IV.1

INITIAL BACKWARD STEPWISE REGRESSIONS

Model	1	2	3	4	5	6	7
N	288	288	288	288	288	288	288
F	14.75	16.03	17.51	19.24	21.29	23.80	26.66
R^2	.41164879	.41156857	.41098725	.40985281	.40806474	.40557905	.39990753
SSE	2.50449332	2.50483480	2.50730936	2.51213844	2.51974992	2.53033101	2.55447354
Exogenous Variables				**COEFFICIENTS (t-ratio)**			
Intercept	0.78342304	0.78320041	0.78464413	0.78857740	0.78920620	0.79694134	0.78969866
E	-0.00000393 (-5.9615434)	-0.00000394 (-6.0489668)	-0.00000399 (-6.1854668)	-0.00000402 (-6.2601916)	-0.00000408 (-6.3765194)	-0.00000399 (-6.2904689)	-0.00000398 (-6.2481997)
CL	-0.31410589 (5.4194095	-0.31373742 (-5.4249423)	-0.31915753 (-5.6178287)	-0.31961877 (-5.6311632)	-0.31986163 (-5.66373752)	-0.32345108 (-5.7078892)	-0.31826548 (-5.6089214)
PCH	0.09077408 (4.8826222)	0.09067599 (4.8867166)	0.09166259 (4.9729267	0.08937317 (4.9244280)	0.08895970 (4.9050993)	0.08829316 (4.8692915)	0.08946032 (4.9234134)
COUNTY	-0.07086870 (-2.703011)	-0.07079058 (-2.7055498)	-0.07140491 (-2.7349588)	-0.07060260 (-2.7092434)	-0.07330303 (-2.8319604)	-0.6919655 (-2.7018512)	-0.10510370 (-8.005623)
LOND	-0.08832586 (2.1886068)	-0.09144887 (2.4779023)	-0.09214992 (-2.5019992)	-0.08369466 (-2.3958297)	-0.08253560 (-3.3664319)	-0.07991301 (-2.295648)	-0.07866174 (-2.2538855)
SCOT	0.09779459 (5.2810983)	0.09763027 (5.2877216)	0.09755460 (5.2905576)	0.09758094 (5.296225)	0.09837558 (5.3469617)	0.0982179 (5.376802)	0.09592235 (5.2249401)
WAL	0.04997654 (2.0880613)	0.05003149 (2.0952326)	0.4880362 (2.0566963)	0.04861193 (2.0493901)	0.04916673 (2.0760539)	0.04916805 (2.0736441)	0.04646060 (1.9595917)
AREA	-0.00166183 (-1.421267)	-0.00164609 (-1.4142135)	-0.00167024 (-1.43527)	-0.00170799 (-1.4730919)	-0.00164409 (-1.4177446)	-0.00186167 (-1.6309506)	
ELCHG	0.08548677 (1.161895)	0.08499366 (1.1575836)	0.08781291 (1.2)	0.08288967 (1.1401754)	0.07834879 (1.0816653)		
HOUS	-0.00000002 (-.9110433)	-0.00000002 (-.9110433)	-0.00000002 (-.9219544)	-0.00000002 (-0.9165151)			
PROPAV	0.00007484 (0.7)	0.00007718 (0.7280109)	0.00007684 (0.7280109)				
PET	0.00991406 (0.5196152	0.00996612 (0.5196152)					
DENS	-0.00000003 (-0.2)						

To test for a significant improvement made by a more complex form of the variable we use an F statistic. We will use a more complex form if the improvement is significant at the ten-percent level ($\alpha = .10$), making 2.73 the critical value for "F" with 1 and 290 degrees of freedom.

It is not clear whether turnout should respond to the absolute difference in votes between last winner and first loser (CLM), to the difference as a proportion of eligible voters (CL), or to the difference as a proportion of all votes cast (CLV), or to the difference as a proportion of the votes for the last winner and the first loser (CLN), assuming no plumping, The exploratory regressions we ran over these variables are presented in Table IV.2. We have 291 degrees of freedom and the SSE's are:

	SSE		F
CLM	2.600364		
		>	.5577
CL	2.59538		
		>	11.31
CLV	2.499299		
		>	13.48
CLN	2.383496		

Now using CLN as our closeness variable, we look at several competing relationships between turnout and electorate size. The relationship may be linear (E), logarithmic (LE), or polynomial (E, E^2). The exploratory regression over these variables are presented in Table IV.3. The corresponding SSE's are:

TABLE IV.2

SEARCH FOR A BETTER BASE FOR THE CLOSENESS VARIABLE

Model	1	2	3	4
N	300	300	300	300
F	27.728	30.397	33.901	27.594
R^2	0.3993	0.4215	0.4483	0.3981
$\overline{R}^2$	0.3849	0.4077	0.4351	0.3837
SSE	2.595380	2.499299	2.383496	2.600364

Exogenous Variables — COEFFICIENTS (t-ratio) [VIF]

Exogenous Variables	1	2	3	4
Intercept	0.789585 (44.748) [0.000000]	0.789584 (45.710) [0.000000]	0.790265 (46.902) [0.000000]	0.759462 (42.928) [0.000000]
E	-.0000039564 (-6.358) [1.38995]	-0.000041578 (-6.825) [1.383286]	-.000003646 (-6.086) [1.402950]	-.0000012636 (-1.571) [2.317614]
CL	-0.314734 (5.826) [1.233755]			
CLV		-0.270363 (-6.817) [1.237614]		
CLN			-0.285754 (-7.932) [1.194789]	
CLM				-.0000284374 (-5.773) [1.834390]
PCH	0.089320 (4.999) [1.054904]	0.087357 (5.018) [1.039936]	0.092631 (5.431) [1.047007]	0.89565 (5.004) [1.834390]
COUNTY	-0.104690 (-8.231) [1.186540]	0.107810 (-8.637) [1.188391]	-0.108183 (-8.937) [1.171860]	-0.103316 (-8.155) [1.176625]
LOND	-0.079414 (-2.314) [1.404036]	-0.076443 (-2.276) [1.396664]	-0.087496 (-2.659) [1.405697]	-0.101078 (-2.890) [1.455177]
SCOT	0.096014 (5.448) [1.262203]	0.104124 (5.970) [1.283613]	0.102728 (5.149) [1.234892]	0.087072 (5.071) [1.195804]
WAL	0.041338 (1.856) [1.041437]	0.044864 (2.051) [1.043123]	0.040711 (1.908) [1.040839]	0.042473 (1.905) [1.042119]

TABLE IV.3

SEARCH FOR A BETTER ALGEBRAIC FORM FOR THE ELECTORATE SIZE VARIABLE

Model	1	2	3	4
N	300	300	300	300
F	33.901	39.240	30.805	37.021
R^2	0.4483	0.4841	0.4585	0.4702
$\overline{R}^2$	0.4351	0.4717	0.4437	0.4575
SSE	2.383496	2.229037	2.339363	2.289023
Exogenous Variables	COEFFICIENTS (t-ratio) [VIF]			
Intercept	0.790266 (46.902) [0.000000]	1.109396 (23.363) [0.000000]	0.800034 (46.424) [0.000000]	0.833087 (44.481) [0.000000]
E	-0.000003646 (-6.086) [1.402950]		-.0000068503 (-4.599) [8.779355]	
LNE		-0.41640 (-7.735) [1.301679]		
E^2			6.78983E-11 (2.343) [8.002580]	
ESR				-0.000134134 (-7.114) [1.384895]
CLN	-0.285754 (-7.932) [1.194789]	-0.265650 (-7.568) [1.212982]	-0.269708 (-7.410) [1.238637]	-0.270998 (-7.630) [1.209474]
PCH	0.092637 (5.431) [1.047007]	0.087306 (5.281) [1.051667]	0.093681 (5.532) [1.047731]	0.090471 (5.408) [1.048429]
COUNTY	-0.108183 (-8.937) [1.171860]	-0.087160 (-7.112) [1.284547]	-0.098374 (-7.733) [1.314179]	-0.096700 (-7.975) [1.224217]
LOND	-0.087496 (-2.659) [1.405697]	-0.098397 (-3.308) [1.228064]	-0.079675 (-2.427) [1.420384]	-0.081347 (-2.573) [1.350790]
SCOT	0.102728 (6.149) [1.234892]	0.090277 (5.519) [1.266955]	0.092141 (5.362) [1.326624]	0.093286 (5.640) [1.260690]
WAL	0.040711 (1.908) [1.040839]	0.043250 (2.101) [1.035972]	0.037167 (1.751) [1.046153]	0.039195 (1.875) [1.040648]

	SSE		F
E	2.383496		
		>	18.86
Ln(E)	2.229037		
		>	(negative)
E, E^2	2.339363		

We will be using the logarithmic form of the electorate size variable. Once again, we look at alternative algebraic forms for the closeness variable. In Table IV.4 we try a logarithmic form (LnCLN); an exponential form similar to the cumulative normal density function [$\exp(-CLN^2)$]; a polynomial form (CLN and CLN^2) and, of course, the linear form, CLN. The SEE's are:

	SSE		F
CLN	2.229037		
		>	2.92
CLN, CLN^2	2.183975		
		>	(negative)
$\exp(-CLN^2)$	2.185924		
		>	(negative)
Ln(CLN)	2.544		

The logarithmic form has a higher SSE than the linear form and will not be considered further. The polynomial form has a lower SSE than the exponential form and is also much simpler. The "F" for the improvement made by the polynomial form over the linear form is 2.92, which is a significant improvement, so we choose the polynomial form of the term.[22]

Our second step of the search is now complete. We reduced the SSE from 2.600364 to 2.183975, a sixteen-percent reduction in the SSE. We will use the logarithmic form as the electorate-size variable and the polynomial form for the closeness term.

TABLE IV.4

SEARCH FOR A BETTER ALGEBRAIC FORM FOR THE CLOSENESS VARIABLE

Model	1	2	3	4
N	300	300	300	300
F	33.140	40.735	35.585	29.130
R^2	0.4841	0.4941	0.4945	0.4112
$\overline{R}^2$	0.4717	0.4819	0.4806	0.3971
SSE	2.119037	2.185924	2.183975	2.543999

Exogenous Variables	COEFFICIENTS (t-ratio) [VIF]			
Intercept	1.109396 (23.363) [0.000000]	0.571787 (6.655) [0.000000]	1.101397 (23.336) [0.000000]	1.079664 (20.094) [0.000000]
LNE	-0.041640 (-7.735) [1.302679]	-0.041663 (-7.826) [1.299064]	-0.041532 (-7.781) [1.302767]	-0.046710 (-8.211) [1.274695]
CLN	-0.265650 (-7.568) [1.212982]		-0.092917 (-1.182) [6.189030]	
EXP(-CLN^2)		0.527588 (8.010) [1.87948]		
CLN			0.270943 (2.450) [5.956180]	
LNCLN				-0.015491 (-3.746) [1.111985]
PCH	0.087306 (5.281) [1.051667]	0.080941 (4.987) [1.033668]	0.082181 (4.991) [1.067237]	0.079030 (4.444) [1.065940]
COUNTY	-0.087160 (-7.112) [1.284547]	-0.087281 (-7.212) [1.277223]	-0.087429 (-7.194) [1.067237]	-0.079030 (-5.521) [1.23002]
LOND	-0.098397 (-3.308) [1.228064]	-0.100628 (-3.415) [1.229218]	-0.087429 (-7.194) [1.284652]	-0.070742 (-2.323) [1.209514]
SCOT	0.090277 (5.519) [1.266055]	0.091449 (5.658) [1.260626]	0.091377 (5.632) [1.267026]	0.016814 (3.661) [1.172119]
WAL	0.043250 (2.101) [1.035972]	0.039627 (1.944) [1.035595]	0.040949 (2.004) [1.038166]	0.045056 (2.046) [1.039146]

Our third step is to add those variables excluded in the first step, since some may become significant with our improved algebraic forms. Again, we employ a backward stepwise procedure. There are two basic results of this step; first, to add the percentage change in electorate size (ECH) back into the regression model, and second, to delete the "unsquared" closeness variable. The final model is:

$$\text{Turnout} = B_0 + B_1\text{LNE} + B_2\text{CLN}^2 + B_3\text{ECH} + B_4\text{PCH} + B_5\text{COUNTY} + B_6\text{LOND} + B_7\text{SCOT} + B_8\text{WAL} + \varepsilon.$$

The coefficients and appropriate statistics of this model can be seen in Table IV.6. The results of the intermediate steps generated by the stepwise procedure can be seen in Table IV.5.

V. Interpreting the Coefficients

Now that we have selected the variables to be included in our model, we can interpret the coefficients for the 1880 OLS regression (used in our exploratory search) and the 1865, 1868, and 1874 GLS regressions.[23] The GLS regressions will be used in our two tests for structural change in the next chapter. These coefficients can be seen in Table IV.6.

Because of the non-linearities in the variables CLN^2 and LNE, we shall use a different approach in interpreting their coefficients. For each coefficient, we shall look at the effect on turnout at the mean and the extreme values for those variables in that year.

ECH: The coefficient of the electorate change variable means that a one percent increase in the electorate size over the previous election is correlated with a .14895 increase in turnout in 1880; a .08565

TABLE IV.5

FINAL BACKWARD STEPWISE REGRESSIONS

Model	1	2	3	4	5	6	7
N	288	288	288	288	288	288	288
F	20.86	22.54	24.50	26.70	29.26	32.31	35.84
R^2	0.51681537	0.51677188	0.51668202	0.51553352	0.51372998	0.51124399	0.50682449
SSE	2.05682025	2.0570005	2..05738791	2.06227682	2.06995415	2.06995415	2.09934943
Exogenous Variables	COEFFICIENTS (t-ratio)						
Intercept	1.10839553	1.10991337	1.10846644	1.10774463	1.11181022	1.11823731	1.11226463
LNE	-0.04399259 (-7.6452599)	-0.044139 (-7.7834529)	-0.04391770 (-7.8816241)	-0.04469487 (-7.9611556)	-0.04469487 (-8.0727038)	-0.04481785 (-8.0901174)	-0.04506130 (-8.1154174)
CLN^2	0.31300907 (2.8)	0.31643964 (2.8017851)	0.31672396 (2.8089143)	0.39817884 (7.8809897)	0.39783207 (7.87400782)	0.39808140 (7.8733728)	0.39229177 (-7.7575769)
PCH	0.08400272 (4.9558046)	0.08426797 (5.003984)	0.08438047 (5.0219518)	0.08194551 (4.9608467)	0.08146492 (4.9335585)	0.07807956 ((4.7968739	0.07935883 (4.8682645)
COUNTY	-0.05119717 (-2.1330729	-0.5129184 (-2.1424285)	-0.05147073 (-2.1540659)	-0.05061648 (-2.1213203)	-0.05313174 (-2.2405356)	-0.05192306 (-2.1886968)	-0.08398083 (-6.7520367
LOND	-0.11756312 (-3.3867388)	-0.1177189 ((3.6701498)	-0.11439349 (-3.6701498)	-0.11381001 (-3.6551333)	-0.11314273 (-3.6345562)	0.10176687 (-3.4322004)	-0.10099614 (-3.3970575)
SCOT	0.09069360 (5.2591044)	0.09058500 (5.36656)	0.09083511 (5.4018515)	0.08911557 (5.3460265	0.08977429 (5.3898051)	0.8981705 (5.3879495)	0.08702771 (5.235456)
WAL	0.04808584 (2.2181073)	0.047754689 (2.2181073)	0.4769576 (2.2181073)	0.04721500 (2.197726)	0.04784233 (2.2293496)	0.04769089 (2.2203603)	0.04533970 (2.1095023)
ELCHG	0.13875389 (2.0420577)	0.13991975 (2.068816)	0.13991975 (2.0736441)	0.14188822 (2.1071307)	0.13776705 (2.0493901)	0.13022454 (1.9442222)	0.14895332 (2.2538855)
AREA	-0.00166403 (-1.5652475)	-0.00166942 (-1.5748015)	-0.00168707 (-1.5968719)	-0.00168000 (-1.5905973)	-0.00161404 (-1.5329709)	-0.00166980 (-1.5842979)	
PROPAV	0.001138303 (1.1789826)	0.00011384 (0.0049874)	0.00011121 (1.1661903)	0.00011393 (1.195826)	0.00011345 (1.1916375)		
HOUS	-0.00000002 (-0.9899494)	0.000000002 (0.9949874)	-0.00000002 (-1.00)	0.00000002 (1.0148891)			
CLN	-0.06257451 (-.7745967)	-0.06438137 (-.8062258	-0.06458061 (-.8062258)				
DENS	0.00273965 (0.7071067)	0.00000003 (0.2236)					
PET	0.00273965 (0.1414213)						

TABLE IV.6

GLS REGRESSIONS FOR THE 1865, 1868, AND 1874 ELECTIONS
AND OLS REGRESSION FOR THE 1880 ELECTION

Model	1865	1868	1874	1880
N	170	131	241	245
F	27.536	51.044	42.379	36.347
R^2	0.5777	0.7700	0.5937	.5045
$\overline{R}^2$	0.5568	0.7549	0.5979	.4906
SSE	0.234834	0.097439	0.130365	2.079624

Exogenous Variables — COEFFICIENTS (t-ratio) [VIF]

Exogenous Variables	1865	1868	1874	1880
Intercept	1.016468 (20.525) [0.000000]	0.864650 (20.145) [0.000000]	1.346879 (24.476) [0.000000]	1.109215 (23.726) [0.000]
LNE	-0.032906 (-4.763) [1.650748]	-0.034157 (-6.936) [1.683880]	-0.068946 (-10.147) [1.900624]	-0.44634 (-8.316) [1.319427]
CLN^2	-0.462013 (-10.000) [1.132343]	-0.328340 (-3.443) [1.454986]	-0.194415 (-3.732) [2.389468]	-0.38332 (-0.946) [1.167689]
ELCHG	-0.00446228 (-0.089) [1.246285]	0.118447 (2.217) [7.339303]	0.085697 (1.598) [3.654745	0.145722 (2.305) [1.062588]
PCH	0.0005475899 (0.034) [1.225185]	-0.989565 (12.278) [1.761945]	0.005833723 (0.516) [1.533836]	0.078819 (4.93) [1.030427]
COUNTY	-0.076723 (-3.960) [2.322192]	-0.064555 (-3.156) [17.27292]	-0.05575 (-6.678) [4.307116]	-0.086383 (-7.178) [1.268254]
LOND	-0.284959 (-2.317) [1.035423]	-0.060689 (-0.888) [1.131799]	-0.185179 (-0.056) [1.000022]	-0.092772 (-3.084) [1.202378]
SCOT	0.062850 (3.305) [1.412817]	0.166830 (12.407) [7.456323]	0.024708 (2.784) [2.294692]	0.08956 (5.63) [1.253644]
WAL	0.065675 (2.218) [1.133155]	-0.027503 (-0.961) [1.051532]	-0.045749 (-3.206) [1.423719]	0.043256 (2.148) [1.30990]

increase in 1874; a .11845 increase in 1868; but a .0045 decrease in 1865 (where it is insignificantly different from zero at the ten percent confidence level). The coefficient for 1874 is only significant at the twelve percent confidence level. Before the regression was run, we were unsure of the direction of the effect. New members of the electorate in modern U.S. elections vote less often than members who have had the vote for a long time and have the "voting habit." In Britain in this period, voting may have given the individual greater social status. Being able to vote meant being available to receive bribes, and the low-income citizens (to whom the franchise was extended in 1867) were usually more receptive to bribes. Of course, new voters, since they were interested enough to register, probably voted more often then those who had long had the franchise, and since marginal turnout increases, average turnout should increase.

<u>PCH</u>: The coefficient of the binary variable, <u>party change</u>, is positive in all four elections, but significant only in the 1868 and 1880 elections (.098956 and .07093588 in 1868 and 1880, respectively). This means that a difference in the party affiliation of the last winner and first loser is associated with about a ten percentage point increase in turnout in the 1868 election and about an eight percentage point increase in the election of 1880).

<u>COUNTY</u>: The coefficients for the binary variable, <u>county</u>, are all significantly negative, as expected. The values range from -.064555 in 1868 to -.085575 in 1874. This means that county districts tend to have a turnout which is between 6 1/2 and 8 1/2 percentage points smaller than in the borough districts, other things equal.

LOND: The London binary variable has negative coefficients for all four of the elections, but only in the 1868 and 1880 elections are these coefficients significantly different from zero. The coefficient for this variable is -.28495 in 1865; -.0607 in 1868; -.1852 in 1874; and -.101 in 1880. The London districts had turnouts that, on average, were twenty-eight percentage points lower in 1865; six percentage points lower in 1868; eighteen percentage points lower in 1874; and ten percentage points lower in 1880.

SCOT: The binary variable, Scotland, has an estimated coefficient of .06285 for the 1865 election; .16683 for 1868; .024798 for 1874; and .08703 for 1880. This means that the Scottish districts had voter turnout which was from two percentage points higher in 1874 to sixteen percentage points higher in 1868, other things equal.

WAL: The binary variable, Wales, seems to have an ambiguous effect, with coefficients of 0.065675 in 1865, -0.27503 in 1868; -.04575 in 1874; and .04534 in 1880. Two of these have significantly positive coefficients and one has a significantly negative coefficient. The Welsh districts had significantly larger turnout by 6 1/2 percentage points in 1865 any by 4 1/2 percentage points in 1880, but significantly lower turnout in 1874 by 4 1/2 percentage points.

CLN^2: Our closeness variable, is the squared percentage difference between the votes received by the last winner and votes received by the first loser, based on the sum of the votes those two candidates received. The range for this coefficient was -0.462 in 1865 to -0.194 in 1974. In 1880 and 1868 the coefficients were -0.392 and

-.328, respectively. The 1874 coefficient seems to be "out of line" with the others.

Because of the non-linear nature of this variable, we shall interpret the coefficients at the mean and maximum values for the variable. At the means of CLN^2, compared to what turnout would have been if elections had been as close as possible turnout was reduced by .0437 in 1880; by .007834 in 1874; by .0121 in 1868; and by .018713 in 1865, ceteris paribus. At the largest values of CLN^2 in the respective years, turnout was reduced by .36017 in 1880; by .136785 in 1874; by .3164 in 1868; and by .462013 in 1865.

LNE: This variable is also non-linear; it is the natural log of the electorate size. The coefficients for this variable are -.0329 in 1865; -.034157 in 1868; -.068946 in 1874; and -.04506 in 1880. Comparing the effect in the districts with the smallest electorate size in each election year with that year's mean electorate size, we see that turnout in the smallest districts was higher than in the districts of average size by 8.74 percentage points in 1880, by 12.83 percentage points in 1874, by 6.2 percentage points in 1868, and by 6.5888 percentage points in 1865, ceteris paribus. Comparing the effect in the districts with the largest electorate size with that year's mean electorate size, we see that turnout was smaller than in the districts of average size by 10.755 percentage points in 1880, by 18 percentage points in 1874, by 8.07 percentage points in 1868, and by 9.157 percentage points in 1865, ceteris paribus.

VI. Summary

In this chapter we developed an empirical model of voter participation based on the calculus of voting. We rigorously developed a weighting procedure to correct for heteroscedasticity. We then went through a multi-step exploratory search to separate the variables that have detectable effects from those that do not, honing the algebraic forms of the quantitative variables. We then looked at the regression results, interpreting the coefficients.

In the next chapter we use these regression results along with regressions on some pooled data to test for structural changes between elections. Keep in mind that the data set used in the exploratory search is not used in testing hypotheses in this next chapter. To do otherwise would be to destroy any validity that these tests have.

FOOTNOTES

[1]Op. cit., Judge, p. 586.

[2]Loc. cit.

[3]Loc. cit,

[4]Ibid, p. 587.

[5]Loc. cit.

[6]Arnold Zellner and T.H. Lee, "Joint Estimation of Relationships Involving Discrete Random Variables," Econometrica, Vol. 33, 1965, p. 387.

[7]Op. cit., Judge, p. 587.

[8]T. Toyoda, "Use of the Chow Test Under Heteroscedesticity," Econometrica, Vol. 42, #3, May, 1974, pp. 601-608.

[9]Op. cit., Ferejohn and Fiorina, 175, p. 922.

[10]Our population figures are inhabited houses, and property values are from 1880. The population figures for north and south Lanarkshire are approximations only, since the figures in McCallmont's was only for the entire county, not the Northern and Southern county parlimentary districts separately, so that the populations in the two districts was assumed to be proportional to the number of inhabitant houses in the two districts.

[11]Op. cit., Craig, p. xiv.

[12]Op. cit., Craig, p. xvi.

[13]We also may change the functional forms of some of the following variables in our search procedure, but transformations will be positive monotonic, keeping the original flavor of the variables we now discuss.

[14.5]Op. cit., Gash, p. 117.

[15]J. Vincent and M. Stenton, ed., McCalmont's Parlimentary Poll Book, 8th ed., 1971, Harvester Press, Ltd., Brighton, Sussex, passim.

[16]Op. cit., Hanham, Elections and Party Management, 1978, p. 160.

[17]Ibid, p. 244

[18]Op. cit., Judge, et al., pp. 410-417.

[19]Jane T. Helwig and Kathryn A. Council, SAS User's Guide, Cary, North Carolina: SAS Institute Inc., 1979.

[20]They define relative efficiency as

$$R_0 = \frac{M_1}{M_2} = \frac{n/\sum_{obs}(mse_i(y^*) \text{ for forward selection})}{n/\sum_{obs}(mse_i(y^*) \text{ for backward deletion})}$$

$$= \frac{\sum_{obs}(mse_i(y^*) \text{ for backward deletion})}{\sum_{obs}(mse_i(y^*) \text{ for forward selection}},$$

where y^* is the estimator for the true value of y for any case. Kennedy and Bancroft, "Model Building for Prediction Based on Repeated Significance Tests," Annals of Mathematical Statistics, Vol. 42, pp. 1273-1284.

[21]To detect multicollinearity, we report the variance inflation factor (VIF) in the tables in this chapter and the next, for each variable. The Variance Inflation Factor (VIF) is given in Table II to detect multicollinearity problems. If the regressor variables have been standardized by the length of the column vector, than the k^{th} diagonal element of (X'X) will be $\frac{1}{1-r^2_{kk}}$, where r^2_{kk} is defined as the coefficient of determination of a regression of the k^{th} variable on all other regressors. Since the sampling variance of the k^{th} element of β, σ^2_{kk}, is $\sigma^2(\frac{1}{1-r^2_{kk}})$, multicollinearity will cause σ^2 to be inflated (as $r^2_{kk} \to 1$, $\sigma^2_{kk} \to \infty$). The k^{th} diagonal element is given the name, Variance Inflation Factor (VIF). Op. cit., Judge, pp. 461-462.

[22]For the polynomial form F must have degrees of freedom of 2 and 189, respectively, which has a critical value of 2.30.

[23]One observation from the borough district, Wigan, England, was excluded from our data set because the number of votes cast for candidates was greater than twice the number of electors, though each elector had two votes to cast. See the returns for the 1868 election in Craig, 332.

CHAPTER 5

THE EFFECTS OF THE BALLOT ACT AND THE SECOND REFORM ACT ON TURNOUT

In this chapter we use the empirical model developed in the previous chapter to test two sets of hypotheses. In the next section we test for some structural change between last the election before the Secret Ballot Act and the first election after it. Following that, we test a number of hypotheses concerning the changes that the literature suggests would occur because of the Second Reform Act greatly extending the franchise.

I: A Test of the Cox-Kousser Hypothesis

Given the weak and indirect statistical evidence in Cox and Kousser that the introduction of the secret ballot will cause a significant change in voter turnout, we wish to test the null hypothesis that the electorates before and after the secret ballot are drawn from the same population. To test this hypothesis, a linear regression model corrected for heteroscedasticity, as discussed in the previous chapter, is estimated for the elections immediately before and after the Ballot Act of 1872; the 1868 election and the 1874 election. These election data sets are then pooled, and the model is estimated for both elections. Kmenta gives the appropriate F-test for pooling data as:

$$F_{K,n+m-2K} = \frac{(SSE_c - SSE_1 - SSE_2)/K}{SSE_1 + SSE_2/(n+m - 2K)}$$

where: SSE_1 and SSE_2 are the sum of squared errors from the 1868 and 1874 election models, SSE_c is the sum of squared errors from the pooled regresion model; n is the number of observations in the 1868 election; m is the number of observations in the 1874 election; and K is the number of regressors including the intercept.[1]

Looking at Table V.2, we can see that our calculated value of F is greater then the critical value. This test leads us to reject the hypothesis that the two data sets were drawn from the same population; there was a significant change. The questions now are, "Can we locate the variable or variables in which this change occurred?" and "Could the secret ballot have caused these changes in the electorates' responses to these variables?" To locate the source of this structural shift, a dummy-interaction model is used, with the dummy variable equal to one if the election occurred in 1874 and zero if it occurred in 1868. If the year-dummy variable has a significant t value, there was a change in the intercept term. If the year-interaction variable has a significant t value, there was a change in the eletorate's response to changes in that variable.

As can be seen in Table V.1, the significant change occurred in the intercept term and the coefficients for the electorate size, the party-change and the Scottish variables. The significant increase in the intercept means that after correcting for all of the factors we could, turnout actually increased in 1874 instead of decreasing. This runs counter to the Cox-Kousser hypothesis. The effect of electorate size was significantly larger in 1874 than in 1868. The difference-in-party variable seems to have had a substantially weaker effect and the Scottish turnout was not as far ahead of the British

TABLE V.1

POOLED AND DUMMY-INTERACTION REGRESIONS

Model	Pooled 1865/1868	Pooled 1868/1874	Pooled 1865/1868 DUMMY-INTERACTION		Pooled 1868/1874 DUMMY-INTERACTION	
N	301	373	301		373	
F	48.544	34.258	33.666		47.516	
R^2	0.5708	0.4295	0.6691		0.6947	
$\overline{R}^2$	0.5591	0.4170	0.6493		0.6801	
SSE	0.431014	0.425661	0.332272		0.227805	
Exogenous Variables				DUMMY INTERACTION		DUMMY INTERACTION
Intercept	0.960532 (27.960) [0.000000]	0.988355 (26.311) [0.000000]	1.016468 (22.876) [0.000000]	-0.151818 (-2.219) [130.072196]	0.864650 (22.475) [0.000000]	0.482203 (6.862) [803.568676]
LNE	-0.033100 (-7.043) [1.918845]	-0.037869 (-8.772) [1.298969]	-0.329062 (-5.309) [4.195397]	-0.00125116 (-0.145) [142.253605]	-0.034157 (-7.738) [2.474645]	-0.034785 (-4.094) [827.558702]
CLN^2	-0.420741 (-10.188) [1.042837]	0.045133 (0.815) [1.08608]	-0.462013 (-11.146) [1.320825]	0.133673 (1.088) [2.763237]	-0.328340 (-3.841) [7.100490)	0.133953 (1.313) [12.046715]
ELCHG	-0.010266 (-0.539) [1.491597]	0.033331 (1.879) [3.999990]	-0.00446228 (-0.099) [10.561161]	0.122910 (1.556) [36.994534]	0.118447 (2.473) [53.116333]	-0.032849 (0.6602) [8.258134]
PCH	0.060627 (7.630) [1.8209086]	0.026419 (4.484) [3.156171]	0.0005475894 (0.038) [7.572978]	0.098408 (5.659) [5.678217]	0.098956 (13.697) [8.649487]	-0.093122 (-6.611) [31.929748]
COUNTY	0.085600 (-6.566) [5.580504]	-0.076715 (-7.602) [4.199011]	-0.076723 (-4.414) [12.437058]	0.012168 (0.402) [43.775318]	-0.064555 (-3.520) [25.262447]	-0.021027 (-0.919) [82.725956]
LOND	-0.150501 (-2.074) [0.0390]	-0.128536 (-1.634) [1.026242]	-0.284959 (-2.582) [3.045935]	0.224269 (1.625) [3.141449]	-0.060689 (-0.0991) [1.132479]	-0.124497 (-0.035) [1.000325]
SCOT	0.133136 (11.347) [4.907373]	0.092702 (10.205) [3.504638]	0.062850 (3.684) [13.045152]	0.103980 (4.406) [26.564172]	0.166830 (13.842) [11.243805]	-0.142121 (-9.267) [36.978349]
WAL	0.038723 (1.689) [1.071878]	-0.00188947 (-0.113) [0.9098]	0.026569 (2.472) [2.808865]	-0.093178 (-2.133) [1.669512]	-0.027503 (-1.073) [5.112660]	-0.018249 (-0.612) [5.521385]

TABLE V.2

F-STATISTICS FOR POOLING TESTS

SSE

SSE_{65} = .234834

SSE_{68} = .097439

SSE_{74} = .130366

$SSE_{65/68}$ = .431014

$SSE_{68/74}$ = .425661

Pooled Elections	Degrees of Freedom Numerator	Degrees of Freedom Denominator	Level of Significance	Critical Value for F	Calculated Value for F
65/68	8	283	.95	1.94	10.51233
68/74	8	354	.95	4.389	34.16317

turnout in 1874 as it was in 1868. Although significant at only the .20 level, it seems as if the response of turnout to closeness weakened in 1874 compared to 1868. There is a problem with multicollinearity, as can be seen by the large VIF's (variance inflation factors) for many of the variables, which can cause significant variables to appear to be insignificant, by inflating the variance.

The question now is "Could a change in the bribery have caused a decreased response of turnout to the closeness of the election?" The answer is yes, but only indirectly. Before the secret ballot, a close race could be won, by buying enough votes.[3] Candidates are more likely to purchase votes when the election is close, that is, the "price" of a seat is lower when the projected outcome is closer, increasing turnout. After the secret ballot, a close race could be tilted by paying supporters of your opponent to stay home on election day, tending to decrease turnout.[4] This decrease in turnout could be more than offset, however, by voters' increased participation in close elections, but the response of turnout to closeness is still weakened. This is the Denver and Hand's explantion of the effects of closeness on voter turnout.[5]

This suggests that a simultaneous-equations model may be more appropriate.[6] If so, what we have estimated would be a reduced-form equation. This does not invalidate our results regarding a structural change, because the closeness coefficient can only change if either voters' reactions to closeness change or if candidates' responses to closeness change, and there is no reason that we can see for the voters' responses to closeness to have changed over this period.

To support the Cox-Kousser position requires great imagination, and even then the evidence for that support is weak. Running counter to their hypothesis is the strong evidence that, after controlling for the effects of other variables, turnout was increased with the introduction of the secret ballot. Changes other than the Ballot Act could have caused the structural shift detected by the Chow test above. This bit of evidence from England changes slightly the case for the Cox-Kousser position enough to suggest investigation for alternative causes of the decline in turnout in the U.S.

II: The Extension of the Franchise

> As for popular suffrage, it may be further remarked that especially in large states it leads inevitably to electoral indifference, since the casting of a single vote is of no significance where there is a multitude of electors. Even if a voting qualification is highly valued and esteemed by those who are entitled to it, they still do not enter the polling booth. Thus, the result of an institution of this kind is more likely to be the opposite of what was intended; election actually falls into the power of a few, of a caucus, and so of the particular and contingent interest which is precisely what was to have been neutralized.[6]
>
> G.W.F.Hegel

The introduction of the secret ballot was not the only interesting change in the British electoral history in the days of Gladstone and Disraeli. Extensions and restrictions of suffrage have been of primary concern to those political scientists using empirical methods to study popular participation in elections[7] and the extension of the franchise under the Second Reform Act in 1867 can be investigated.

The negative coefficient for electorate size in Table V.6 makes us expect that the effect of an extension of the franchise on voter turnout is negative. However, some of the positive coefficients for the electorate change variables lead us to the opposite position.

Settle and Abrams, in their time-series estimation of voter turnout (measured as a percentage of voter age-eligible population), use a dummy variable to control for the impact of women's suffrage.[8] Though doubling the number of those fulfilling the requirements to vote, the extension of the franchise to women under the nineteenth amendment increased participation in elections of those over 21 by only eighteen percent.[9]

Chambers and Davis[10] suggest a reason that might account for a reduction in turnout, as we measure it, from an extension of the franchise. The newly enfranchised voters are inevitably from lower income groups, and, as Verba and Nie,[11] and Frey[12] point out, lower income individuals tend to vote less often than higher-income individuals. Also, newly enfranchised individuals are not in the habit of voting, but over, time, may become voters. As mentioned in the previous chapter, new voters may be more likely to vote if voting is closely connected with social status, or if they can get paid to vote.

These explanations, however, cannot account for the results of Settle and Abrams.[13] The newly enfranchised women in 1919 were of no lower income than the men of the day (except for single women). The habitual-voter explanation fails to account for the smaller increase in voting by adults than the increase in those qualified to vote, because the time-series was long enough for women to get into the voting habit.[14] We shall see what, if any, light our data analysis can shed on this issue. We have several questions about turnout over the extension of the franchise. First, did the extension of the franchise significantly alter the nature of turnout? Second, if so, what were the factors that were different in their influence on turnout and how were they different? Third, did turnout in the county constituencies change from the franchise extension in a way that was different from the way that turnout changed in the boroughs because of the difference in the franchise extension in the boroughs and the counties? Fourth, did new voters decrease turnout because new voters were yet to be socilaized into voting? Fifth, did the extension of the franchise to the lower income voters decrease turnout? These questions can be addressed using our regression model.

To answer these questions we need to form null hypotheses and develop statistical tests of those hypotheses. Our first hypothesis is that the nature of turnout was not altered by the extension of the franchise. To test this hypothesis, we employ the same sort of before-and-after Chow test as we used in the previous section.

As we can see from Table V.2 our calculated F is greater than the critical value for F, leading us to reject our null hypothesis. We see that there was a change in the nature of turnout.

Our second hypthesis is that the response of turnout to each of the variables was the same before and after the franchise extension. We know that this hypothesis will be rejected because our first hypothesis was rejected. However, we are interested in which variables the change in the structure of turnout occurred. To test this hypothesis, we use a t-test of significance for the dummy and the dummy-interaction variables, where the dummy variable is zero for all observations in the 1865 election and one for all observations in the 1868 election. If the t statistic for the year-dummy variable has an absolute value greater than 1.645 then there was a significant change in the intercept, which would mean that, after all other factors were taken into account, turnout either decreased or increased (depending on the sign). If the t statistic for one of the dummy-interaction variables has an absolute value greater than 1.645, then the relationship between that variable and turnout changed over the extension of the franchise, and the null hupothesis, that there was no change in the relationship, can be rejected. Looking at Table V.1, we can see that t values for the dummy variable and the dummy-interaction with the party change, London, Scotland, Wales, and electorate-change variables had absolute values greater than 1.645. We can reject the hypotheses that there was no change in the effcts of these vriables on turnout over the extension of the franchise.

Our third hypothesis is that there was no difference in the way turnout reacted to the county/borrough difference before and after the franchise extension. To test this hypothesis we use a t test of the dummy interaction with the county variable. To reject this hypothesis, the t statistic must have an absolute value greater than 1.645. As we can see, the t statistic for the interaction of the year dummy with the county dummy is 0.042. We can not reject the null hypothesis that there was no difference between the way turnout changed in the boroughs and the counties between the two elections.

Our fourth hypothesis is that the franchise extension will not decrease turnout because new voters have not gotten into the voting habit yet. We test this by examining the t values for the electorate change variable. If these t statistics are less than -1.282 (a one-tailed test) new voters can be said to have decreased turnout. From Table V.1 we can see that the electorate change variable has the hypothesized sign only for the 1865 election, yet is insignificant. For the 1868 election, it is significant but positive. It is positive and significant at the .12 level for the 1874 election. The non-habitual or new-voter argument does not seem to hold in Britain in this time period. New voters apparently voted in greater proportions than did the more habitual voters.

Our final null hypothesis is that the extension of the franchise to lower income voters did not decrease turnout. To test this, we use a t test of significance of the dummy interaction with the electorate change variable, controlling for the otherwise new voters. If the t statistic for the electorate-change-dummy-interaction variable is less than -1.282, then lower income voters decreased turnout. As we can see in Table V.1, this variable is significantly positive, leading us to reject our null hypothesis. One reason for this could be that voting, so long a privilege for the well-to-do, was still perceived by new voters as giving them status in the community, something they desired. Another reason could be that because lower income individuals were easier to bribe, and for the first time being in a position to receive a bribe, took the opportunity.

What we have discovered by our examination of turnout over the extension of the franchise is that there was a change in the structure of turnout which took place in the intercept, the London variable, the Scottish variable, the Welsh variable, and the electorate-change variable. The 1868 election, after correcting for other factors, has turnout that is sixteen percentage points lower, on average. Turnout is ten percent points higher in 1868 than in 1865 from the party-change effect. The London district effect on turnout is twenty-two percentage points higher. The Scottish district effect is nine and one-half percentage points higher and the Welsh effect was ten percentage points lower, in 1868 than in 1865. The percentage change in electorate size effect is associated with fourteen percentage points higher turnout in 1868 than in 1865.

No difference can be detected between the way the turnout in the boroughs and the counties changed over the franchise extension. We also find that, although turnout declined over the franchise extension, after correcting for other factors, it declined less in the districts the larger the percentage increase in electorate size.

FOOTNOTES

[1]Jan Kmenta, Elements of Econometrics, New York: MacMillan Publishing Co., Inc., 1971, p. 373.

[2]Op. cit., Cox and Kousser, pp. 9-12.

[3]Loc. cit.

[4]Op. cit., Denver and Hands, pp. 33-35.

[5]Leslie Seidle and David Miller, "Turnout, Rational Abstention and Campaign Effort," Public Choice, 27, Fall, 1976, pp. 121-126. Their work suggests a simultaneous equations specification, although they do not use such a specification.

[6]James M. Buchanan, "Hegel on the Calculus of Voting," Public Choice, 17, Spring, 1974, pp. 99-101, quoted from G.W.F. Hegel, The Philosophy of Right, translated with notes by T.M. Knox, Vol. 46, Great Books of the Western World, Encyclopedia Britannica, 1952, Chicago, p. 104.

[7]Joel H. Silbey, Allan G. Bogue, and William H. Flanigan, ed., The History of American Electoral Behavior, Princeton University Press, 1978, p. 137.

[8]Op. Cit., Settle and Abrams, 1976, p. 85.

[9]Ibid., pp. 86-87.

[10]William N. Chambers and Philip C. Davis, "Party, Competition, and Mass Participation: The Case of Democratizing Party System, 1824-1852," in Silbey, Bogue, and Flanigan, p. 175-178.

[11]Sidney Verba and Norman H. Nie, Participation in America: Political Democracy and Social Equality, Harper-Row, 1972, pp. 125-137.

[12]Op. cit., Frey, Bruno, (1971), pp, 101-05; and (1972), pp. 101-103.

[13]Op. cit., Settle and Abrams, pp. 86-87.

[14]Op. cit., Settle and Abrams, p. 82.

CHAPTER 6

SUMMARY AND CONCLUSIONS

I. Summary

This study of Victorian elections began with a review of a debate in political science concerning the decline in voter turnout in the late nineteenth century in the United States. Burnham argued that it was due to alienation resulting from the Republican Party being captured by capitalists. Converse, Rusk, and Cox and Kousser argued that it was the undesired by-product of democratizing political reforms, especially the introduction of the secret ballot, which made paid votes difficult to monitor and therefore decreased the demand for paid votes. Cox and Kousser gave evidence for this proposition, showing that with the introduction of the secret ballot in New York, turnout-inflating fraud decreased and turnout-deflating fraud increased. Our purpose has been to see if this also happened with the introduction of the secret ballot in Great Britain in 1872.

After providing some historical background regarding the political institutions of the reform period in Britain, and after reviewing the theoretical and empirical literature on rational abstention, taking note of the lack of generalization of the most generalized of the calculus-of-voting models, we digressed from the main topic to extend the basic rational abstention model to cover multi-winner and multi-vote elections. First, it was pointed out that changing the number of wimmers merely changes the election's focal point, that is, the candidate pair between whom the real contest takes place. Then it was shown that allowing voters to cast more than one vote in an election (but no more than one vote for any one candidate) alters the voter's calculus. Three general propositions were derived:

Proposition 1: A strategy set, call it T, not containing a voter's favorite, which contains fewer candidates than the number allowed, is dominated by strategy set S, which is T augmented by the voter's favorite candidate. That is, if a voter is casting fewer votes than he is allowed, he must be voting for his favorite candidate. The reason for this is that adding the voter's favorite to the strategy set adds positive expected-utility terms to the voter's calculus equation while adding no negtive terms.

Proposition 2: If a voter's favorite candidate is contained in a strategy set, T, whitch does not contain some candidate, d, which is prefered to at least one element in T, strategy T is not dominated by a strategy S, which is set T augmented by candidate d. The reason for this is that, although at least one positive expected utility term is added to the voter's calculus by adding candidate d, because d is preferred to at least one element in T, at least one negative expected utility term is added because the voter's favorite candidate is preferred to candidate d.

Proposition 3: As long as a voter's strategy set, T, containing the largest allowable number of candidates does not contain his least favorite candidate, strategy T can not be dominated by any other strategy which also contains the largest allowable number of candidates. This is becasue the expected utility of changing the ordering between the least favorite and one of the other candidate, h, may be greater than the expected gain by switching candidate g with h. After this digression, we returned to the main topic, the effects of two reforms on turnout.

Based on the theory of the calculus of voting, in the fourth chapter, an empirical model was developed which could test propositions about the effects of changes in institutions on the structure of turnout. It was first noted that regression models which have binary dependent variables violate the assumption of OLS that the errors are homoscedastic. The observations were then grouped by electoral district, since all voters in a given electoral district faced some of the same circumstances, such as the type of district, the number of electors, etc. This, however, did not cure the heteroscedasticity, but merely put it in a form in which we could do something about it, using GLS and more information. Information from regressions on elections in other years was then combined to estimate error terms which were then used to weight the observations, to improve the efficiency of the estimates and to improve the power of the Chow tests employed in Chapter 5.

After introducing the variables suggested by the calculus of voting literature, we employed a multi-step search procedure to find a model which was expected to have the smallest error sum of squares for the regressions used in the empirical tests. The first step was to purge the model of variables which have no significant effect on voter turnout, by using a backward-stepwise regression technique. Then alternative algebraic forms of the non-binary variables were entered into the model to find the "best" form. All variables which were deleted in the first round of the search procedure were re-introduced and the backward-stepwise selection process was repeated.

After the specification of the model was completed, the estimates of the coefficients for the 1865, 1868, and 1874 GLS regression models and the 1880 OLS regression model were interpreted. One finding seemed rather odd. Although a larger the electorate size was associated with a smaller the turnout rate, we also found that a larger percentage increase in the electorate size was associated with larger turnout rates. The explanation for this that was given was that either the social status connected with the act of voting attracted new voters to the polls at greater rates than people that had been voting previously, or else, since the new voters were lower-income individuals and so were more susceptible to bribes, districts which had the largest increase in these voters had the lowest decrease in turnout.

It was also found that turnout was larger in districts in which the loser with the most votes and the winner with the fewest votes were from different parties. This was as expected, since candidates from different parties are likely to take stances on the issues that are farther apart than would candidates from the same party, which will lead to larger utility difference between the candidates (Δ_{jk}) on average.

The geographic variables, with the exception of the county variable, tended to be unstable. The Welsh variable had two significantly positive coefficients and one significantly negative coefficient. The London variable, though consistently negative and significant, varied by more than a factor of four. The Scottish variable was positive and significant throughout, suggesting that greater social cohesion can lead to higher electoral participation. The significantly negative coefficient of the county variable lent support to the proposition that the higher the cost of voting, the lower the turnout.

The closeness variable was negative and significant. This gave support to the hypothesis that closeness counts in elections, and that the expected-benefit version of rational voter participation models has merit.

Using the empirical model that was developed in Chapter 4, we tested the Cox-Kousser hypothesis, that the introduction of the secret ballot lowered turnout, all other factors being equal. A Chow test was first used to see if there was a significant structural change between the elections before and after the introduction of the secret ballot; there was. Then a dummy interaction model was used to see if turnout was significantly lower after the secret ballot was introduced; it was not. Instead, it was significantly higher, giving evidence contrary to the Cox-Kousser hypothesis, supporting casual observations made at that time that

> It did not appear that the mode of taking votes by ballot had the slightest effect in checking bribery. On the

> contrary, while it enabled more voters to take bribes on both sides, it did not, as fas as we can ascertain, render a single person unwilling to bribe for fear of bribery in vain.[1]

There were significant changes in the coefficients for electorate size, Scotland, and party change, and, at the .20 level, a significant change in the coefficient for the closeness variable.

We also tested for a difference in turnout in the elections before and after the extension of the franchise. We sought answers to several questions concerning the franchise extension. First, did the extension of the franchise significantly alter the nature of turnout? It did. Second, what were the factors that changed and how did they change? We found that the intercept decreased, the party-change coefficient increased, the London and Scotland coefficients increased, and the Welsh coefficient decreased. Third, did turnout in the counties change from the franchise extension in a way which was different from the way turnout chnged in the boroughs? A difference could not be found. Was there evidence that turnout decreased because new voters were yet to be socialized into voting or because the franchise was extended to lower income voters.

Conclusions

There are several conclusions to be made. Some conclusions are rather broad an sweeping, while others are quite specific, pointing out some of the issues that are left unresolved.

First, the expected benefit version of rational abstention seems to have great explanatory power. Turnout varies systemati-cally with closeness and electorate size, determinants of the probability of affecting the outcome. Turnout also varies systematically with social pressure, D, as measured by the Scotland variable, and with the costs of voting, C, as measured by the county-borough variable. By comparison, the evidence for minimax regret is quite weak.

Second, although turnout seemed to decrease with the introduction of the secret ballot in the United States, introduction of the secret ballot in Britain was not connected with a decline in turnout, though the Cox-Kousser hypothesis suggests that turnout should have declined. There must have been some difference in the way candidates and their agents in the U.S. and Britain responded to the secret ballot. For some reason, candidates did not cut back on paying electors to vote in Britain. Why this is so remains an open question to be answered by future research efforts.

Third, turnout decreased when the franchise was extended by the Second Reform Act, both in counties and boroughs, in ways which were not different, in spite of the difference between the way the franchise was increased in the counties (by 45%), and in the boroughs (by 134%). This also remains a puzzle.

Last, we see that changes in political institutions, the rules of the game, can alter turnout. This change in turnout can alter the election outcome if the increase in votes on one side is not cancelled out by an equal increase on the other side. These changes in election outcomes can cause differences in policy, which can, in democratic countries where a rule change only requires a simple majority, alter the institutions further, making both rules and outcomes unstable.

FOOTNOTES

[1]H.J. Hanham, The 19th Century Constitution: Documents & Commentary, Cambridge at the University Press, London, 1969, p. 290, from the Report of the Royal Commission of Enquiry concerning corrupt practices at Sandwich in 1880.

BIBLIOGRAPHY

Aldrich, John J., "Some Problems in Testing Two Rational Models of Participation," American Journal of Political Science, 20, November, 1976, pp. 713-733.

Aranson, Peter H., Melvin Hinich and Peter Ordeshook, "Elections Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," American Political Science Review, March, 1974, pp. 135-152.

Beck, Nathaniel, "A Note on the Probability of a Tied Election," Public Choice, Vol. 23, The Hague: Martinus Nijhoff Publishers, Fall, 1975.

__________, "The Paraxod of Minimax Regret," American Political Science Review, 60, September, 1975, p. 918.

Black, Jerome, "The Multicandidate Calculus of Voting: Application to Canadian Federal Elections," American Journal of Political Science, 20, August, 1978, pp. 609-638.

Buchanan, James M., "Hegel on the Calculus of Voting," Public Choice, 17, Spring, 1974, pp. 99-101.

Burnham, Walter Dean, "Theory of Voting Research: Some Reflections on Converse's 'Change in the American Electorate'," American Political Science Review, 68, 1974.

__________, "The Changing Shape of the American Political Universe," American Political Science Review, 59, March, 1965, pp. 7-28.

Brams, Stephen J. Brams and Peter C. Fishburn, "Approval Voting," American Political Science Review, Vol. 72, 1978.

Barzel, Yaram and Eugene Silverberg, "Is the Act of Voting Rational?" Public Choice, 16, Fall, 1973, p. 51-58.

Brock, Michael, The Great Reform Act, Hutchinson and Co., Ltd., London, 1973, p. 18.

Brody, Richard A. and Benjamin I. Page, "Indifference, Alienation and Rational Decisions," Public Choice, 15, Summer, 1973, pp. 1-17.

Cain, Bruce E. "Strategic Voting in Britain," American Journal of Political Science, 22, August, 1978, pp. 639-655.

Campbell, A., Gerard Gurin, and William E. Miller, The Voter Decides, Westport, Conn.: Greenwood Press, 1971.

Carroll, Lewis, Alice in Wonderland, edited by Donald J. Gray, New York: W.W. Norton and Company, Inc., 1971.

Chamberlain, Gary and Michael Rothschild, "A Note on the Probability of Casting a Decisive Vote," Journal of Economic Theory, 25, August, 1981.

Chambers, William and Philip Davis, "Party, Competition, and Mass Participation: The Case of Democratizing Party System, 1824-1852," in Silbey, Bogue and Flanagan, eds., The History of American Electoral Behavior, 1978.

Coats, R. Morris, "Closeness in Elections," unpublished paper presented to the Department of Political Science, State University of New York at Stoney Brook, 1981.

Converse, Philip E., "Comment on Burnham's 'Theory and Voting Research'," American Political Science Review, 68, September, 1974, pp. 1024-1027.

__________, "The Concept of the Normal Vote," in Campbell, et. al., Elections and the Political Order, New York: John Wiley and Sons, 1966, p. 24.

Cowling, Maurice, 1867, Disraeli, Gladstone and Revolution, Cambridge at the University Press, London, 1967, p. 14.

Cox, Gary W., and J. Morgan Kousser, "Turnout and Rural Corrpution: New York as Test Case," American Journal of Political Science, 25, Nov. 1980, pp. 646-664.

__________, "Turnout and Rural Corruption: New York as a Test Case," Social Science Working Paper 292, Pasadena: California Institute of Technology, October, 1979. The revised form of this paper has been cited in Footnote No. 2.

Craig, F.W.S., ed., British Parliamentary Election Results 1832-1885, London: MacMillan Press, 1977, p. xiv.

Denver, D.T. and H.T.G. Hands, "Marginality and Turnout in British General Elections," British Journal of Political Science, 4, January, 1974, pp. 17-35.

Downs, Anthony, An Economic Theory of Democracy, New York: Harper and Row, 1957.

Ferejohn, John A., and Morris Fiorina, "Closeness Counts Only in Horseshoes and Dancing," American Political Science Review, 69, September, 1975, pp. 920-925.

Fraser, John, "Political Participation and Income Level: An Exchange," Public Choice, 13, 1972, pp. 115-118.

Freund, John, Mathematical Statistics, Englewood Park, New Jersey: Prentice-Hall, 1971.

Frey, Bruno S.,"Why do High Income People Participate more in Politics?" Public Choice, 11, Fall, 1971, pp. 101-105.

Gash, Norman, Politics in the Age of Peel, 2nd ed., Sussex, England: Harvester Press, Ltd., 1977.

Good, I.J. and Lawrence S. Mayer, "Estimating the Efficacy of a Vote," Journal of Behavioral Science, 20, 1975.

Gosnell, Harold F., Getting Out the Vote, Chicago: University of Chicago Press, 1927.

Gywn, William B., Democracy and the Cost of Politics in Britain, London: The Althalone Press, 1962, pp. 61-92.

Hadley, Arthur T., The Empty Polling Booth, Englewood Cliffs: Prentice-Hall, 1978.

Hanham, H.J., Elections and Party Management, Sussex England: Harvester Press, 1978, p. 160.

__________, ed., Dod's Electoral Facts 1832-1853 Impartially Stated, Brighton, Sussex England: Harvester Press, 1972.

Helwig, Jane T., and Kathryn A. Council, SAS User's Guide, Cary, North Carolina: SAS Institute Inc., 1979.

Judge, George, et.al., The Theory and Practice of Econometrics, New York: John Wiley and Sons, 1980.

Key, V.O., Southern Politics in State and Nation, New York: Alfred A. Knopf, 1950.

Kmenta, Jan, Elements of Econometrics, New York: MacMillan Publishing Co., Inc., 1971, p. 373.

Ledyard, John, "The Paradox of Voting and Cnadidate Competition: A General Equilibrium Analysis," California Institute of Technology Social Science Working Paper 224, 1975.

Leamer, Edward E., Specification Searches, New York: John Wiley and Sons, Inc., 1978.

Lloyd, Trevor, The General Election of 1880, London: Oxford University Press, 1968, pp. 24-29.

Luce, R. Duncan and Howard Raiffa, Games and Decisions, New York: John Wiley & Sons, Inc., 1957.

Lawrence, S. Mayer and I.J. Good, "Is Minimax Regret Applicable to Voting Decisions?", American Political Science Review, 69, September, 1975, pp. 916-917.

Margolis, Howard, "The Probability of a Tie Election," Public Choice, 31, Fall, 1977.

McKelvey, Richard D. and Peter Ordeshook, "A General Theory of the Calculus of Voting," Mathematical Applications in Political Science, Vol. 4, James F. Herndon and Joseph L. Bernd (eds.), Charlottesville: The University of Virginia Press, 1972.

Niemi, Richard, "Costs of Voting and Non-voting," Public Choice, Vol. 27, 1976.

Palfrey, Thomas R. and Howard Rosenthal, "Team Games and Voter Participation," Carnegie-Mellon University Graduate School fo Industrial Administrtion Working Paper 25-882, 1982.

Pinto-Duchinsky, Michael, British Political Finance: 1830-1980, American Enterprise Institute, Washington, D.C., 1981.

"Representation of the People Act, 1867," British Parliamentary Sessional Papers, 1867-68, Vol. XX, pp. 1-799.

Riker, William H., and Peter C. Ordeshook, "A Theory of The Calculus of Voting," American Political Science Review, 62, March, 1969.

Rusk, Jerrold G., "Comment: The American Electoral Universe," American Political Science Review, 68, September, 1974, pp. 1041-1042.

Russell, Keith, "Political Participation and Income Level: An Exchange," Public Choice, 13, Fall, 1972, pp. 113-114.

Settle, Russell, and Baron A. Abrams, "The Determinants of Voter Participation: A More General Model,: Public Choice, 27, Fall, 1976, pp. 81-89

Seidle, Leslie, and David Miller, "Turnout, Rational Abstention and Campaign Effort," Public Choice, 27, Fall, 1976, pp. 121-126.

Silbey, Joel H., Allan G. Bogue, and William Flanigan, eds., The History of American Electoral Behavior, Princeton: Princeton University Press, 1978.

Siblerman, Jonathan and Garey Durden, "The Rational Behavior Theory of Voter Participation: The Evidence from Congressional Elections," Public Choice, 23, Fall, 1975, pp. 101-108.

Smith, F.B., The Making of the Second Reform Bill, London: Cambridge at The University Press, 1966.

Strom, Gerald, "On the Apparent Paradox of Participation: A New Proposal," American Political Science Review, 69, September, 1975, pp. 908-913.

Thompson, Fred, "Closeness Counts in Horseshoes and Dancing ..., and Elections," Public Choice, 38, 1982, pp. 305-316.

Tollison, Robert and T.D. Willett, "Some Simple Economics of Voting Not Voting," Public Choice, Fall, 1973, pp. 59-71.

Tollison, Robert, Mark Crain, and Paul Pautler, "Information and Voting: An Empirical Note," Public Choice, 24, Winter, 1975, pp. 43-49.

Tullock, Gordon, Toward A Mathematics of Politics, Ann Arbor: The University of Michigan Press, 1967.

__________, "The Paradox of Not Voting for Oneself," American Political Science Review, 69, September, 1975, p. 919.

Toyoda, T., "Use of the Chow Test Under Heterskedasticity," Econometrica, 42, May, 1974, pp. 601-608.

Verba, Sydney and Norman H. Nie, Participation in America: Political Democracy and Social Equality, 1972, pp. 125-137.

Vincent, J., and M. Stenton, McCalmont's Parliamentary Poll Book, 1832-1918, Brighton, Sussex: Harvester Press, 8th ed., 1971.

Webster's Third New International Dictionary, editor-in-chief, Philip Babcock Gove, Springfield, Mass.: G. & C. Merriman Company, 1961.

Zechman, Martin J., "Dynamic Models of the Voter's Decision Calculus: Incorporation of Retrospective Considerations into Rational-Choice Models of Individual Voting Behavior," Public Choice, 34, 1979, pp. 197-215.

Zellner, Arnold, and T.H. Lee, "Joint Estimation of Relationships Involving Discrete Random Variables," Econometrica, 33, April, 1965, pp. 382-394.

The vita has been removed from the scanned document