

<b>I. Einleitung</b>	<b>3</b>
<b>II. Political Stock Markets (PSM)</b>	<b>5</b>
<b>II.1 Double Auction Markets (DAM)</b>	<b>5</b>
<b>II.2 Geschichte der Political Stock Markets</b>	<b>9</b>
<b>II.2.1 Iowa Political Stock Markets (IPSM)</b>	9
<b>II.2.2 Austrian Political Stock Markets</b>	11
a) APISM94	11
b) APISM95/96	13
<b>II.3 Begriffe und Erläuterungen</b>	<b>18</b>
<b>II.3.1 Information</b>	18
<b>II.3.2 Asymmetrische Information, Informationseffizienz, Motivation</b>	18
<b>II.3.3 Arbitrage</b>	19
<b>II.3.4 Marginal Traders</b>	21
<b>II.3.5 News-Days</b>	22
<b>II.4 AEM/APISM97</b>	<b>23</b>
<b>II.4.1 Design des APISM97</b>	23
<b>II.4.2 Teilnehmeranalyse</b>	29
<b>II.4.3 Händlerverhalten</b>	32
<b>II.4.4 Ergebnisse und Auswertungen der Märkte</b>	36
a) Allgemeine Resultate	36
b) ÖVP-Aktie	38
c) SPÖ-Aktie	39
d) FPÖ-Aktie	41
e) Grüne-Aktie	42
f) LiF-Aktie	44
<b>II.5 Zusammenfassung</b>	<b>46</b>
<b>III. Prognosen durch Meinungsumfragen</b>	<b>48</b>
<b>III.1 Umfrageforschung in der empirischen Sozialforschung</b>	<b>48</b>
<b>III.1.1 Methoden der empirischen Sozialforschung</b>	48
<b>III.1.2 Prognosen und Hypothesenbildung</b>	49
<b>III.1.3 Trenduntersuchungen</b>	50
<b>III.2 Die Stichprobenauswahl</b>	<b>51</b>
<b>III.2.1 Grundgesamtheit</b>	51
<b>III.2.2 Statistische Grundkenntnisse</b>	53
<b>III.2.3 Strukturierung des Objektbereiches</b>	56
a) Objektbereich, Merkmal & Variable	56
b) Klassifikation	57
c) Typologie	59
<b>III.2.4 Auswahlverfahren</b>	59
a) Die reine Zufallsauswahl und das systematische Sample	61
b) Stratifizierte Stichproben	62
c) Mehrstufiges Auswahlverfahren und Clusterstichprobe	64
d) Die Flächenstichprobe (Random-Route-Verfahren)	66

e) Das Quotenverfahren	67
<b>III.2.5 Nichterfassungsfehler, Ausfallsfehler und Gewichtung</b>	<b>68</b>
<b>III.3 Das Interview in der Umfrageforschung</b>	<b>72</b>
<b>III.3.1 Die Frage</b>	<b>73</b>
<b>III.3.2 Meßqualitätskriterien: Reliabilität und Validität</b>	<b>73</b>
<b>III.3.3 Formulierung von Fragen</b>	<b>78</b>
a) Einleitung	78
b) Fragen, Fragestrategien und Antwortvorgaben	79
c) Aufbau des Fragebogens und Positionierung einer Frage	82
<b>III.3.4 Interviewer-Befragten-Beziehung</b>	<b>83</b>
<b>III.3.5 Das Telefoninterview</b>	<b>86</b>
<b>III.4 Umfragen zur Oberösterreichischen Landtagswahl</b>	<b>87</b>
<b>III.4.1 Umfragen zu politischen Wahlen</b>	<b>87</b>
<b>III.4.2 Ergebnisse der Institute bei Prognosen für die OÖLTW</b>	<b>89</b>
a) Strukturierung des Objektbereichs	90
b) Aufbau der Fragebögen und Frageformulierungen	92
c) Ergebniss und Hypothesen	94
d) Weitere Fragestellungen	99
<b>IV. Vergleich der beiden Methoden</b>	<b>103</b>
<b>IV.1 Theoretischer Hintergrund</b>	<b>103</b>
<b>IV.2 Kostenvergleich</b>	<b>104</b>
<b>IV.3 Ergebnisse</b>	<b>106</b>
<b>IV.4 Zukünftige Perspektiven</b>	<b>110</b>
<b>IV.4.1 Prognosen für politische Wahlen</b>	<b>110</b>
<b>IV.4.2 Prognosen von Kennzahlen und in Industrie</b>	<b>111</b>
<b>V. Conclusio</b>	<b>115</b>
<b>VI. Tabellen und Abbildungsverzeichnis</b>	<b>118</b>
<b>VII. Literatur- und Quellenverzeichnis</b>	<b>120</b>

## I. Einleitung

Prognosen vor politischen Wahlen sind von bedeutendem Interesse. Sie haben in den letzten Jahrzehnten die Vorgehensweise der Parteien stark beeinflusst, da bei wesentlichen Sachthemen die Meinung der Bevölkerung naturgemäß für politische Gruppierungen wichtig und notwendig ist.

Vor allem die Medien schenken Umfrageergebnissen große Aufmerksamkeit, und man kann schon fast postulieren, daß dadurch die Bevölkerung die Themen für die Politik vorgibt und nicht umgekehrt. Beispiele dafür, daß Politiker auf Grund von Meinungsumfragen „umgefallen sind“, gibt es zur Genüge.

Ziel dieser Arbeit ist es, die Möglichkeiten, Wahlprognosen zu erstellen, anhand der Oberösterreichischen Landtagswahl 1997 zu analysieren und zu vergleichen, und zwar:

- Prognosen durch repräsentative Meinungsumfragen und
- Prognosen durch Bewertung durch Börsen (Political Stock Markets)

Die klassische Methode der demoskopischen Prognose wird von fast allen Meinungsforschungsinstituten praktiziert. Diese Methode weist schon eine lange Geschichte auf, die Theorie ist sehr weit fortgeschritten. Es wird versucht, durch die Befragung einer kleinen, repräsentativen Gruppe (Stichprobe) auf die gesamten Wähler mittels statistischer Stichprobentheorie „hochzurechnen“. Es ist klar, daß sich dadurch nur Prozentsätze mit gewissen Schwankungsbereichen angeben lassen können, die von der Größe der Stichprobe abhängig sind.

Anders funktionieren Political Stock Markets: Die Repräsentativität der Teilnehmer ist nicht gefordert, diese handeln mit Aktien der Parteien mit den anderen Teilnehmern. Grundlage der Transaktionen sind aber nicht die Präferenz für bestimmte Parteien, sondern nur die Annahmen, wie diese bei der Wahl abschneiden werden.

Die Methodik der beiden Verfahren ist also grundsätzlich unterschiedlich. Die Frage ist nun: Welche Methode liefert die „besseren“ Ergebnisse? Welche Potentiale und möglichen Anwendungsgebiete liegen vor?

Die Arbeit wird die prinzipiellen Vorgehensweisen erklären und dem Leser verständlich machen, und die obigen Fragen anhand der Oberösterreichischen Landtagswahl zu klären versuchen.

In Kapitel II wird das Prinzip der Political Stock Markets erläutert. Dort werden Double Auction Markets, die Geschichte der Political Stock Markets (PSM), einige Charakteristika der PSM und konkrete Ergebnisse des APSM97 anhand der Oberösterreichischen Landtagswahl (OÖLTW) vorgestellt.

Kapitel III beschäftigt sich mit der Methodik der Wahlprognosen durch Meinungsumfragen. Es sollen die unterschiedlichen Ansätze bezüglich Stichprobenauswahl, Einflüsse der Befragungsart und der Fragestellung vorgestellt werden.

Statistische Auswertungen der Umfrageergebnisse, konkrete Ergebnisse der OÖLTW verschiedener Institute und zusätzliche Fragestellungsmöglichkeiten, um das Stimmungsbild der Bevölkerung auszuloten, werden weiters behandelt.

Kapitel IV vergleicht die beiden Methoden in puncto Treffsicherheit, Risiken, Kosten. Vor allem werden zukünftige Potentiale und Anwendungsmöglichkeiten vorgestellt.

## II. Political Stock Markets (PSM)

### II.1 Double Auction Markets (DAM)

Das Grundkonzept eines Double Auction Market ist sehr einfach. Käufer und Verkäufer stehen sich gegenüber, mit den jeweilig notwendigen Informationen über das gehandelte Gut ausgestattet. Während die Informationen der Verkäufer die Produktionskosten, oder besser gesagt die entstandenen Gesamtkosten zur Bereitstellung des Gutes, sind, sind Käufer naturgemäß mit den Informationen über den Nutzen in Geldeinheiten, der durch den Erwerb eines Gutes generiert wird, ausgestattet.

Käufer und Verkäufer treffen nun auf einem Handelsplatz einander und die jeweiligen Ankaufs- und Verkaufsangebote (*bids* und *asks*) werden ausgetauscht.

Die Charakteristika eines Double Auction Market sind:<sup>1</sup>

- Bids vom Käufer und Asks vom Verkäufer werden öffentlich kundgetan.
- Nach Preisen sortierte Listen werden gebildet, um die Übersicht zu erhöhen.
- Jeder Käufer/Verkäufer kann jederzeit ein Angebot/Gebot akzeptieren.
- Oftmals gilt die Regel, daß neue Angebote/Gebote nur besser sein dürfen wie schon existierende.
- Die Händler haben keine Informationen über Kosten und Nutzen der anderen Händler.
- In den meisten Designs wird der Markt nach einer Periode geschlossen und die Händler können ihre Gewinne kalkulieren.
- In den meisten Designs öffnet der Markt, so wie auch im folgenden Experiment, wieder mit den selben Kosten und Nutzen der Händler. Dieser Vorgang wird oftmals wiederholt.
- Es werden Gleichgewichtspreise und Mengen generiert.

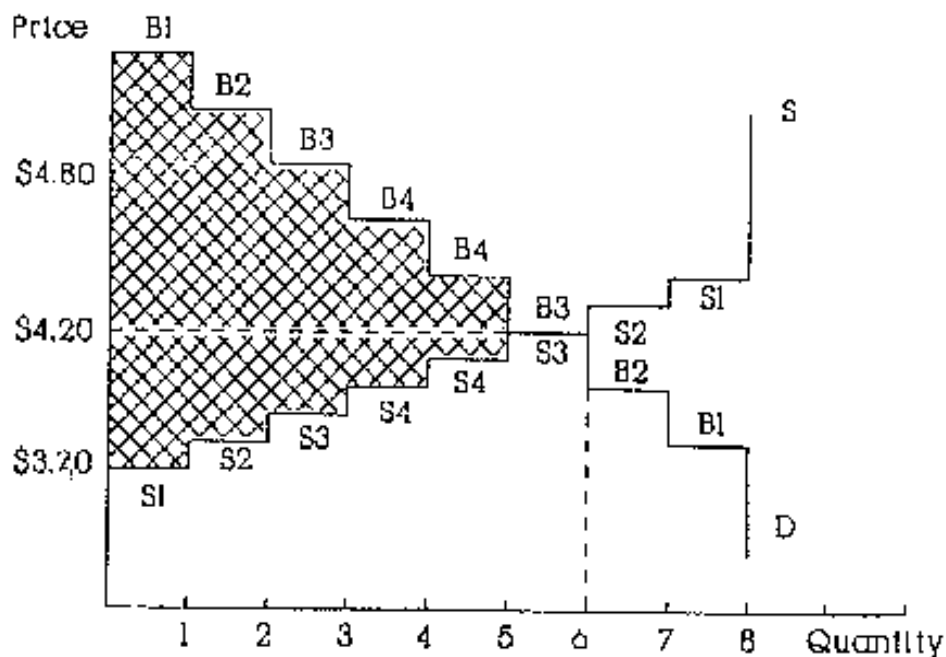
Die letzte Behauptung ist durch zahlreiche Experimente von *Vernon L. Smith*, der als Vater der DAM-Experimente gilt, determiniert. Die Absicht in seinen Experimenten war, eine Umgebung zu schaffen, die die New Yorker Börse simulieren sollte.

Ablauf eines Experiments.<sup>2</sup>

Vor dem Beginn der Perioden werden die Käufer mit den Nutzen und die Verkäufer mit den Kosten der einzelnen Einheiten vertraut gemacht.

Dadurch wird die Angebots- und die Nachfragekurve induziert, die in Abbildung 2.1 dargestellt sind. In diesem Fall schneiden sich die Kurven bei \$4.20 und 5 bzw. 6 Einheiten.

Abbildung 2.1: Induzierte Angebots- und Nachfragekurven



Eine Tradingperiode dauerte 300 Sekunden. Danach konnten Käufer und Verkäufer ihre „Gewinne“ in der abgelaufenen Periode ermitteln und die nächste Periode begann.

---

<sup>1</sup> Vgl. [Davis, Holt 1993] S. 126-130

Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Tabelle 2.1 zeigt den Ablauf der fünften und letzten Periode, Abbildung 2.2 die Ergebnisse des Experiments.

Tabelle 2.1: Verlauf der Abschlüsse in der fünften Periode

Time	Bid	Offer	Time	Bid	Offer	Time	Bid	Offer
296		S2 5.00	241	B2 4.23		164		S2 4.40
294	B4 4.10		237		S4 4.35	161	B4 4.25	
293		S3 4.50	230		S3 4.34	160		S4 4.35
291	B1 4.20		209	B1 4.25		151		S2 4.34
285	B2 4.21		208		S4 4.31	143		S4 4.32
284		S1 4.40	205	B2 4.26		135		S2 4.31
279		S3 4.39	203		S2 accepts	131		S3 4.30
276		S4 4.35	198	B4 4.20		121		S4 4.29
271	B4 4.25		194	B1 4.22		118	B1 4.26	
267	B2 4.26		190		S3 4.40	112		S3 accepts
265		S3 4.34	188	B4 4.25		64		S3 4.28
261		S1 4.30	180		S4 4.35	15	B4 4.25	
254	B4 accepts		176	B3 accepts		13		S4 accepts
249	B2 4.20		171		S1 4.45	10		S2 4.30
245		S3 4.39	167	B4 4.20		0		
244	B1 4.22		165	B1 4.22				

Abbildung 2.2: Ergebnisse des Experiments

<sup>2</sup> Vgl. [Williams 1980] S. 235-258 in [Davis, Holt 1993] S. 130

Am Ende der Periode waren die Preise der gehandelten Güter immer nahe beim Wettbewerbspreis. Die Konvergenz zum Wettbewerbspreis im Laufe des Experiments ist charakteristisch für einen DAM.

In Abbildung 2.2 ist weiters die „Efficiency“ als Gütemaß angegeben. Effizienz in diesem Sinne bedeutet, inwieweit die maximal zur Verfügung stehende Konsumenten- und Produzentenrente (schraffierter Bereich in Abbildung 2.1) ausgenutzt werden konnte.

$$\text{Effizienz } E = \frac{\sum(MV_i - P_i) + \sum(MC_i - P_i)}{\varepsilon_C} \times 100\%$$

$\varepsilon_C$ ...Maximal erzielbare Produzenten- und Konsumentenrente

MV...Nutzen einer Einheit in Geldeinheiten

MC...Kosten einer Einheit

P.....Preis einer Einheit

Die Effizienz liegt schon in der vierten Periode bei 100%!

Weiterführende Varianten dieses Grundkonzepts, die zum schnelleren Erreichen von Gleichgewichtspreisen führen können, sind in weiteren Experimenten<sup>3</sup> getestet worden:

- nur jeweils eine Order (bid/ask) pro Trader
- nur Verbesserung bezüglich Preis oder Menge führt zu neuen Orders
- nur eine Einheit pro Order

Das einfache System von DAM hat in die unterschiedlichsten Märkte Einzug gehalten, vom Fließhandel mit Aktien und Wertpapieren bis zu Warentermingeschäften baut eine große Vielfalt an Märkten auf diesem Konzept auf.

Die beeindruckenden Ergebnisse eines DAM prädestinieren in auch zur Verwendung bei Political Stock Markets. Durch Einsatz von computerisierten DAMs wird kontinuierliches Handeln rund um die Uhr ermöglicht. (*continuous double auction market*).



Computerisierte DAM haben eine Reihe weiterer positiver Effekte, und zwar:<sup>4</sup>

- Kein falsches Auflisten von Angeboten/Geboten.
- Kein inkorrekten Transaktionen
- Standardisierte Präsentation von Instruktionen
- Geringere Preisvolatilität
- Einschränkung von unerlaubter Kommunikation
- Testen von elektronischen Börsen im Feldexperiment

## **II.2 Geschichte der Political Stock Markets**

### **II.2.1 Iowa Political Stock Markets (IPSM)**

1988 wurde an der Universität von Iowa erstmals ein PSM für die damaligen US-Präsidentenwahlen eingerichtet. Es wurden Aktien auf George Bush, Michael Dukakis, Jesse Jackson und „Andere“ auf einem als *continuous double auction market*<sup>5</sup> modellierten Sekundärmarkt gehandelt.

An diesem Experiment nahmen 192 Personen teil, größtenteils naturgemäß Studenten. Weiters ergab eine Untersuchung unter den Teilnehmern, daß 93% weiß und 71% männlich waren. Trotz der „Nicht-Repräsentativität“ der Teilnehmer waren die Ergebnisse äußerst beeindruckend:

- IPSM88-Vorhersage: 53.2%Bush, 45.2%Dukakis
- Wahlergebnis: 53.2%Bush, 45.4%Dukakis

---

<sup>3</sup> vgl. [Smith 1991a] S. 106ff in [Ortner 1996] S. 17

<sup>4</sup> vgl. [Davis, Holt 1993] S. 135-138

<sup>5</sup> Da die Marktabwicklung voll computerunterstützt erfolgt, kann auf den Märkten rund um die Uhr – ohne feste Marktöffnungszeiten – gehandelt werden.

Die Initiatoren des IPSM (Robert Forsythe, Forrest Nelson, George Neumann und Jack Wright) definierten die Unterschiede zu Meinungsumfragen:<sup>6</sup>

- Marktteilnehmer haben einen finanziellen Erfolg vor Augen. Sie haben somit den Anreiz, nach bestem Wissen zu handeln.
- Die Teilnehmer handeln nicht auf der Basis ihrer Präferenzen, sondern auf der Basis ihrer Einschätzung des Ergebnisses.
- Der Markt reagiert schnell auf neue Informationen. Dies ist natürlich wesentlich billiger als eine neue Meinungsumfrage durchzuführen. Ex-post-Analysen lassen sich auf Basis der Kursverläufe durchführen.

Zahlreiche Schlußfolgerungen des Experiments (z.B.: Einfluß der Marginal Traders<sup>7</sup>, Meinungsumfragen sind keine „News“<sup>8</sup>,..) wurden in späteren Experimenten weiterverfolgt.

Dieser Aktienmarkt wird immer als Paradebeispiel hervorgehoben, da die Abweichungen vom tatsächlichen Ergebnis minimal waren, und wesentlich besser als alle von Zeitungen und Meinungsforschungsinstituten durchgeführten Befragungen. Die guten Ergebnisse des ersten IPSM wurden auch bei der Bush-Clinton-Perot Wahl bestätigt.

Zahlreiche weitere Wahl- und Prognosebörsen wurden seitdem eingerichtet:

- PSMs in den USA anlässlich von Präsidentschaft-, Vor-, Senats- und Gouverneurswahlen.
- Die Bonner, Leipziger und Frankfurter Börse anlässlich der Bundestagswahlen 1990 in Deutschland.
- Die Passauer Wahlbörse und Regensburger Wahlbörse 1994 anlässlich der Bundestagswahlen 1994 in Deutschland.
- Der „UBC Election Stock Market“ 1993 und 1996 in Kanada.

---

<sup>6</sup> vgl. [Forsythe, Nelson, Neumann, Whright 1988] S. 2

<sup>7</sup> vgl. auch II.3.4

<sup>8</sup> vgl. auch II.3.5

- Und PSMs in Dänemark, den Niederlanden, Türkei, Finnland und **Österreich.**

### **II.2.2 Austrian Political Stock Markets**

Das erste Experiment dieser Art in Österreich wurde vor der Nationalratswahl 1994<sup>9</sup> an der Technischen Universität Wien am Institut für Betriebswissenschaften, Arbeitswissenschaften und Betriebswirtschaftslehre, Abteilung für Industrielle Betriebswirtschaftslehre (Univ.Prof.Dr.A.Stepan) und Universität Wien, Abteilung Banken und Finanzwirtschaft (Univ.Prof.Dr.J.Zechner) als Austrian Political Stock Market implementiert.

Obwohl beträchtliche Unterschiede bezüglich der Bedeutung und Benützung von Börsen zwischen Österreich und der USA bestehen, wurden die Märkte schnell von den Teilnehmern angenommen und verstanden – dies spricht für das unkomplizierte Design eines *continuous double auction market*.

#### a) APSM94

Am österreichischen Markt für die Nationalratswahlen 1994 wurden Aktienbündel (basic portfolios), die sich aus je einem Stück aller Aktien des Marktes (jeweils 1Stück SPÖ, ÖVP, FPÖ, GRÜNE, LF und Andere) zusammensetzten, am Primärmarkt zum Fixpreis von 10 ATS gehandelt. Die Teilnehmer, die zum Großteil Studenten waren, handelten dann die einzelnen Aktien untereinander am Sekundärmarkt.

---

<sup>9</sup> Neben dem Nationalratswahlmarkt gab es noch den EU-Wahlmarkt, der als Testmarkt diente, und den Koalitionsmarkt. Bei weiteren Betrachtungen wird nur mehr auf den Nationalratswahlmarkt eingegangen.

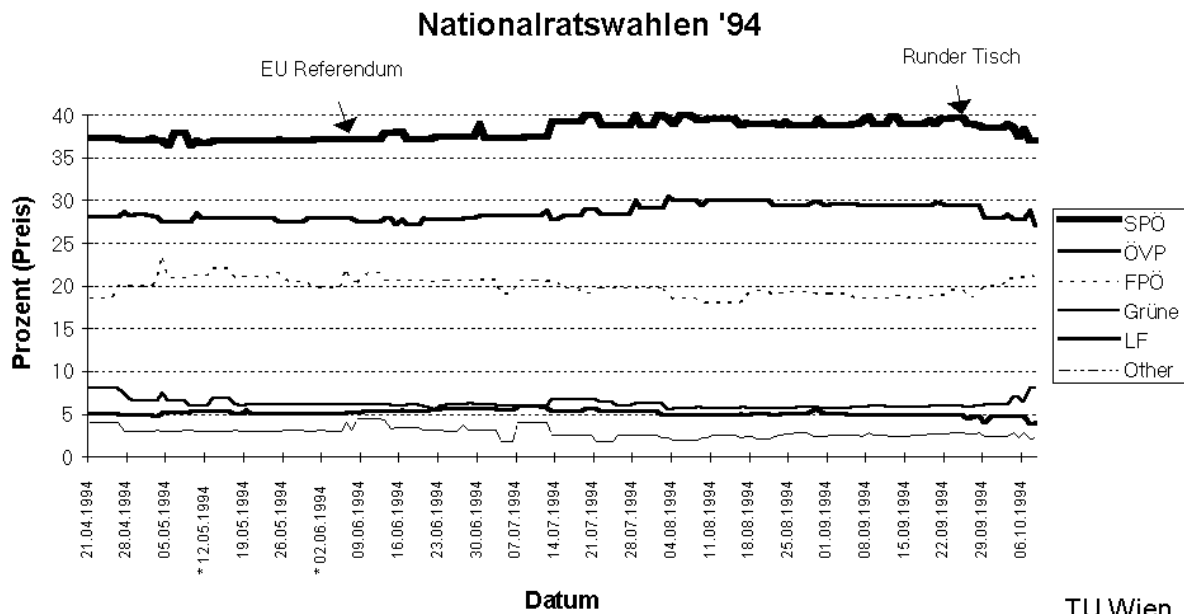
Obwohl die Teilnehmerzahl gering war, fiel das Ergebnis, wie Tabelle 2.2 zeigt, sehr gut aus. Die Kursverläufe sind in Abbildung 2.3 dargestellt, hier sieht man den Einfluß von „News“<sup>10</sup> (EU-Referendum, Runder Tisch im ORF) auf die gehandelten Preise.

Tabelle 2.2: APSM versus Meinungsforschung – Nationalratswahl 1994

Partei	Ergebnis	OGM	d	APSM	d
SPÖ	34,92	38-	3,08	37	2,08
ÖVP	27,67	-26	1,67	27,2	0,47
FPÖ	22,5	22	0,5	21,2	1,3
Grüne	7,31	8-	0,69	8	0,69
LIF	5,97	4 bis 5	1,47	4	1,97
Andere	1,63	-	-	2,2	0,57
TE			7,41		7,08

TE...Total Error;  $TE = \sum d_i$

Abbildung 2.3: Kursverlauf des Nationalratwahlmarkt 94



<sup>10</sup> Vgl. auch II.3.5

b) APSM95/96

Auf Basis der 1994 gewonnenen Erfahrungen wurde eine Verbesserung des Ergebnisses angestrebt und auch erwartet, da der APSM95/96 nicht mehr wie 1994 lokal im Netzwerk der TU Wien, sondern für jedermann erreichbar (im Internet) angeboten wurde. Auf diese Weise wurde:<sup>11</sup>

- a) Ein großer Teilnehmerkreis angesprochen (Alle Österreicher mit Internetanschluß).
- b) Die Bildung einer extrem homogenen Teilnehmergruppe, wie dies im APSM94 der Fall war, verhindert.
- c) In diversen Medien (z.B.: wöchentliche Abbildung des momentanen Kurses im „Profil“) Interesse an dem Projekt geweckt, was zu zusätzlicher Motivation und größerem Teilnehmerkreis führen sollte.
- d) Die nicht über den Markt laufende Kommunikation zwischen Marktteilnehmern stark reduziert bzw. verhindert, da die Teilnehmer sich nur mehr anonym gegenüberstanden.
- e) Weiteres wurde die Beschränkung des erlaubten Investments gelockert (maximal 5.000 ATS).

Trotzdem waren die Ergebnisse wesentlich schlechter wie ein Jahr zuvor. Das Spektrum der möglichen Ursachen ist weitläufig (von Marktmanipulation, zu geringen Teilnehmerzahl bis zu Designfehler).<sup>12</sup>

Tabelle 2.3: APSM versus Meinungsforschung – Nationalratswahl 1995

Partei	Ergebnis	OGM	d	APSM	d
SPÖ	38,06	31,5	6,56	30,66	7,4
ÖVP	28,29	29,5	1,21	27,52	0,77
FPÖ	21,89	25,5	3,61	26,66	4,77
Grüne	4,81	6,5	1,69	7,31	2,5
LIF	5,51	5,5	0,01	6,13	0,62
Andere	1,44	1,5	0,06	1,66	0,22
TE			13,14		16,28

---

<sup>11</sup> Vgl. [Ortner 1996] S. 27,28

<sup>12</sup> Vgl. [Ortner 1996] S. 33-37

# Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Abbildung 2.4: Kursverläufe des Nationratwahlmarkts 95

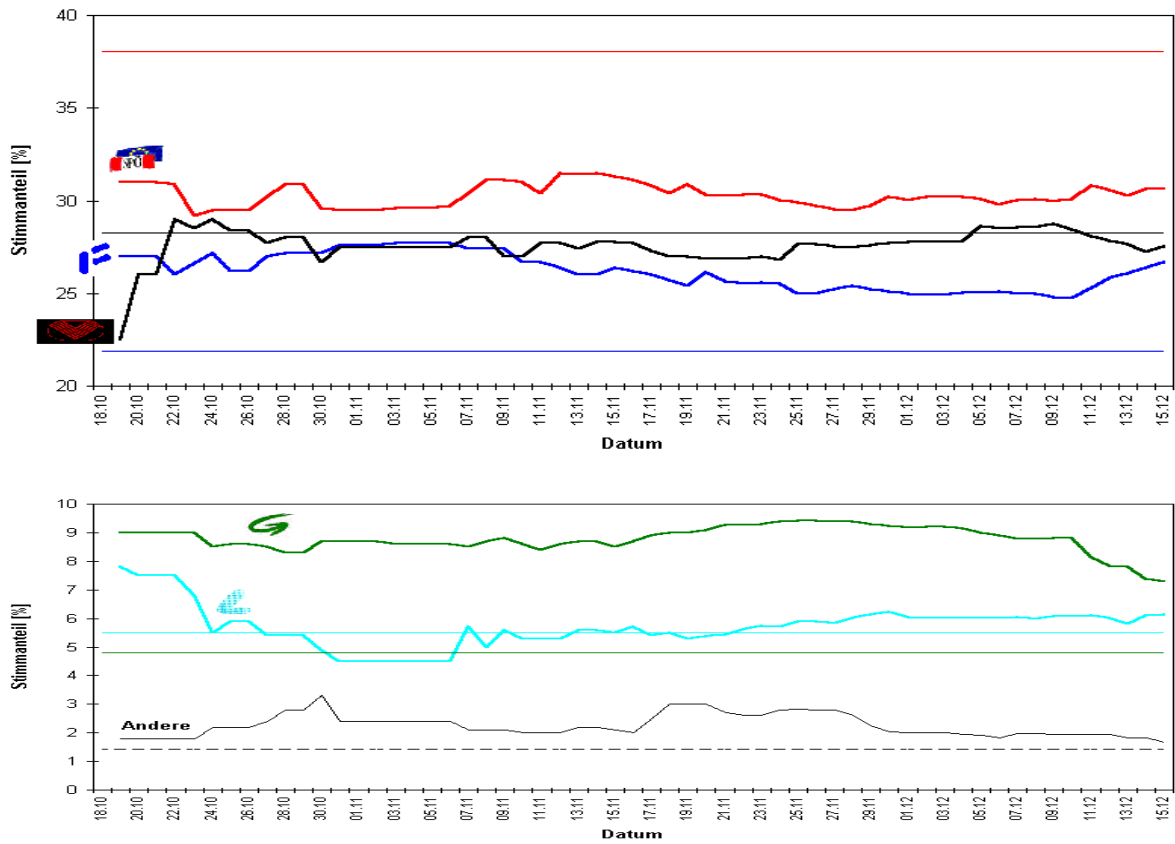
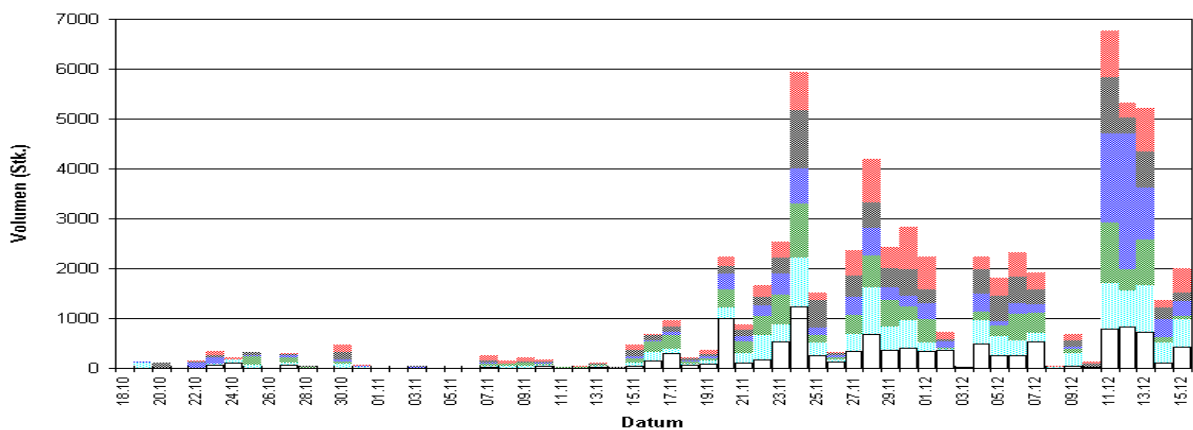


Abbildung 2.5: Handelsvolumen im Nationalratwahlmarkt 95



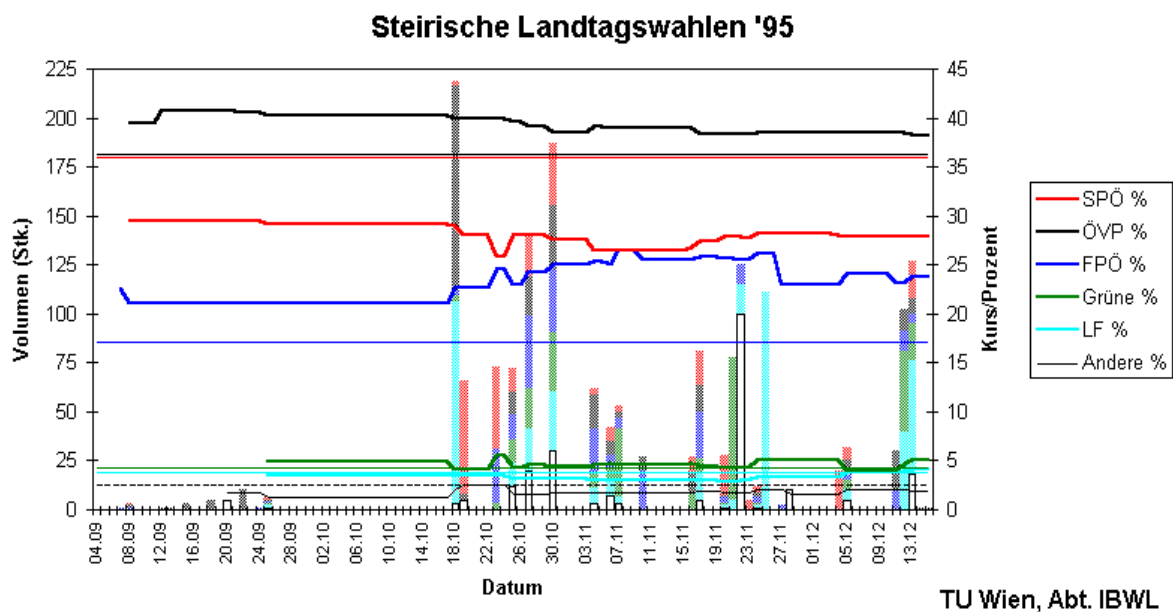
## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Es lief auch ein Markt für die Steirische Landtagswahl, die am selben Tag wie die Nationalratswahl durchgeführt wurde. Das Interesse an diesem Markt war aber äußerst gering, die Ergebnisse entsprechend.

Tabelle 2.4: APSM versus Meinungsforschung – Steirischer Landtagswahlmarkt 95

Partei	Ergebnis	OGM	d	APSM	d
SPÖ	36,04	30,5	5,54	27,91	8,13
ÖVP	36,26	39,5	3,24	38,48	2,22
FPÖ	17,17	19	1,83	23,31	6,14
Grüne	4,23	6	1,77	4,71	0,48
LIF	3,78	4	0,22	4,05	0,27
Andere	2,52	1	1,52	1,94	0,58
TE			14,12		17,82

Abbildung 2.6: Kursverläufe und Handelsvolumen beim Steirischen Landtagswahlmarkt 95



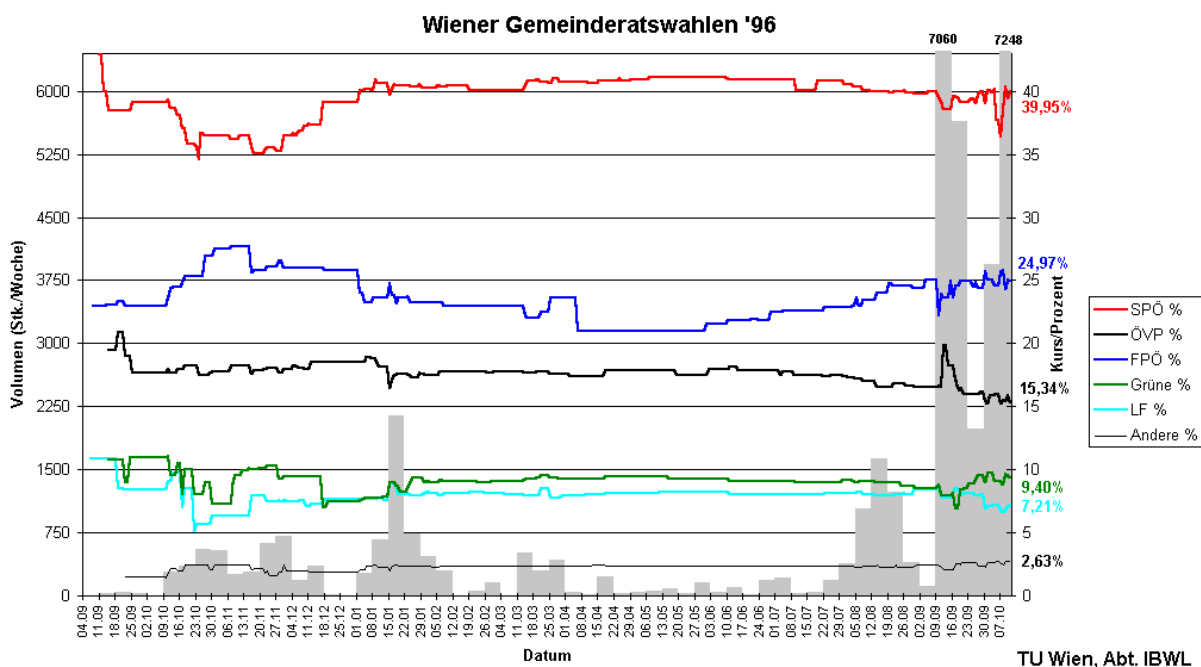
Im Jahr 1996 wurden dann noch zwei APSM durchgeführt und zwar für die Wiener Gemeinderatswahl und die Wahl zum Europäischen Parlament, mit wiederum sehr unterschiedlichen Ergebnissen.

# Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Tabelle 2.5: APSM versus Meinungsforschung – Wiener Gemeinderatwahlmarkt 96

Partei	Ergebnis	Gallup	d	APSM	d
SPÖ	39,15	39	0,15	39,95	0,8
ÖVP	15,26	16	0,74	15,34	0,08
FPÖ	27,94	26	1,94	24,79	2,97
Grüne	7,94	11	3,06	9,4	1,46
LIF	7,95	7	0,95	7,21	0,74
Andere	1,76	1	0,76	2,63	0,87
TE			7,6		6,92

Abbildung 2.7: Kursverläufe und Handelsvolumen beim Wiener Gemeinderatwahlmarkt 95



Während also bei der Wiener Wahl die Ergebnisse des APSM besser waren wie die beste Meinungsumfrage von Gallup, zeigte sich bei der Wahl zum Europäischen Parlament neuerlich der negative Einfluß von geringer Marktteilnahme und Handelsvolumen (siehe Tabelle 2.6 und Abbildung 2.8)

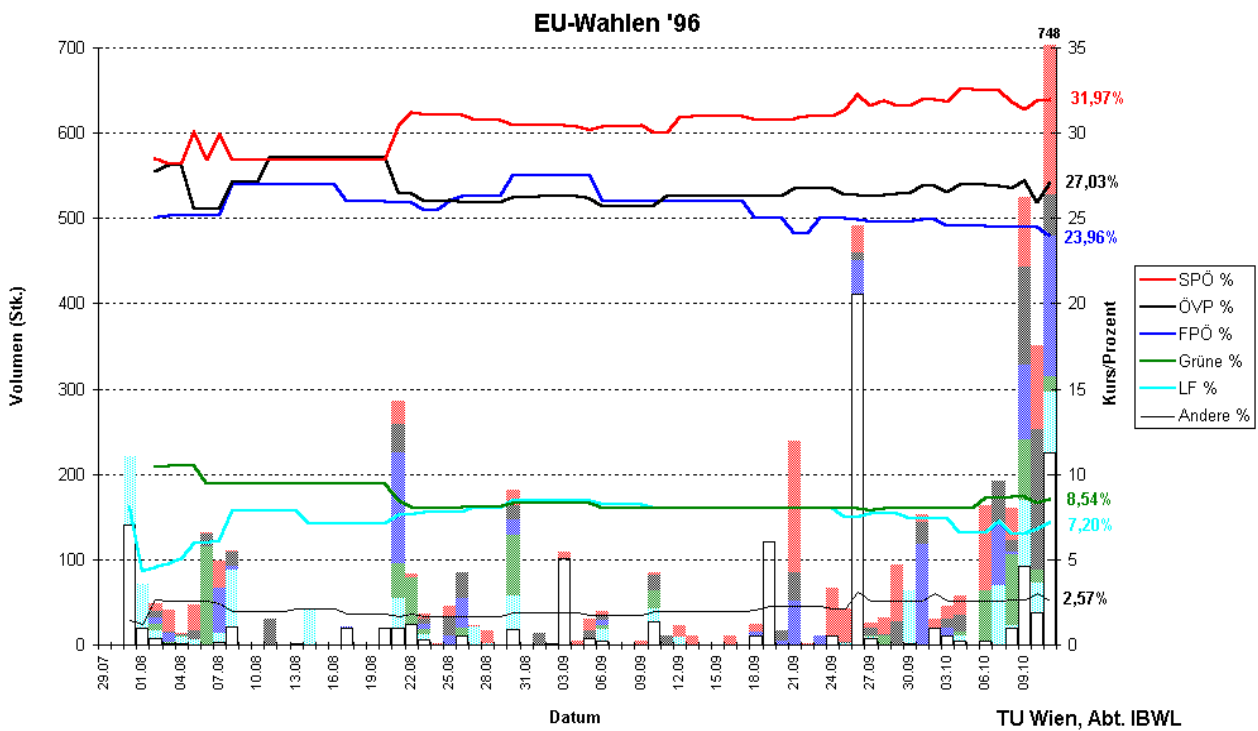


# Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Tabelle 2.6: APSM versus Meinungsforschung – EU-Wahlmarkt 96

Partei	Ergebnis	Gallup	d	APSM	d
SPÖ	29,15	29	0,15	31,97	2,82
ÖVP	29,65	27	2,65	27,03	2,62
FPÖ	27,53	27	0,53	23,96	3,57
Grüne	6,81	10	3,19	8,54	1,73
LIF	4,26	7	2,74	7,2	2,94
Andere	2,6	0	2,6	2,57	0,03
TE			11,86		13,71

Abbildung 2.8: Kursverläufe und Handelsvolumen beim EU-Wahlmarkt



## **II.3 Begriffe und Erläuterungen**

### **II.3.1 Information**

Jeder Marktteilnehmer hat individuelle Vorstellungen (Informationen) über den Ausgang der Wahl. Der Glaube eines Teilnehmers, durch seinen Informationsvorsprung (woher immer er den haben will) oder durch seine bessere Prognosefähigkeit, Renditen am PSM umsetzen zu können, ist die treibende Kraft für den PSM.

Die Motivation ist somit dieselbe wie auf Aktien- oder Optionsmärkten. Es bleibt den Tradern überlassen, sich mit öffentlichen, und somit allen zugänglichen Informationen (z.B.: Meinungsumfragen) zufriedenzugeben, oder sich durch Beschaffung anderwertiger Informationen, die den anderen Tradern eventuell nicht zugänglich sind, Wettbewerbsvorteile zu verschaffen.

### **II.3.2 Asymmetrische Information, Informationseffizienz, Motivation**

Die individuellen Vorstellungen über den Ausgang der Wahl sind notwendig, um Transaktionen überhaupt möglich zu machen. Wenn jeder Teilnehmer wüßte, wie stark eine Partei abschneiden wird, dann gebe es keinerlei Handel. Nur ein unterschiedlicher Informationsstand sichert Transaktionen.

Der Begriff der Informationseffizienz ist eng mit der Kapitalmarkteffizienz verbunden.<sup>13</sup> Ist der Markt 100% effizient, dann fließen der alte Preis und sämtliche neue Informationen in die Bildung des neuen Preises. Der Preis spiegelt also jederzeit den Informationsstand der Marktteilnehmer wider.

Es wird aber schnell die Schwäche der Annahme eines vollkommenen Kapitalmarkts klar, da diese auch impliziert, daß keine Informationskosten entstehen, die

---

<sup>13</sup> vgl. [Miller, Modigliani 1958] S. 261-297

Beschaffung von Information aber immer mit gewissen Aufwänden und Kosten verbunden ist.

So wird ein Teilnehmer am Kapitalmarkt oder am PSM sich nur Information beschaffen, wenn die Kosten der Beschaffung kleiner sind als die dadurch möglichen zusätzlichen erzielbaren Gewinne.

Qualitätsverschlechternde Faktoren des Ergebnisse eines PSM resultierend daraus sind z.B.:

1) Die Anreize für ein gutes Abschneiden ist für den einzelnen Marktteilnehmer zu gering, das Interesse an möglicherweise aufwendiger Informationsbeschaffung ist zu wenig vorhanden.

2) Ein Marktteilnehmer hat gar kein Interesse an einem guten Abschneiden, sondern verfolgt andere Ziele (z.B.: Kursanhebung der FPÖ-Aktie bei der Wahl95 durch einige Marktteilnehmer), handelt deshalb ökonomisch irrational und manipuliert den Markt.

Möglichkeiten, dem entgegenzusteuern, werden in Kapitel II.5 noch genauer besprochen.

### **II.3.3 Arbitrage**

Der Begriff der Arbitrage wird in unterschiedlichen Märkten verschieden verstanden. Auf Wertpapiermärkten versteht man unter Arbitrageure Händler, die versuchen, mit einer rationalen Preisbildung unvereinbaren Kursdifferenzen entweder zwischen zwei vergleichbaren Kontrakten (z.B.: zwei Anleihen unterschiedlicher Laufzeit) oder zwischen zwei gleichen Objekten auf unterschiedlichen Märkten (z.B.: zwischen Future- und Kassamärkten) aufzuspüren und profitabel zu nutzen.

Durch ihre Aktivitäten sorgen sie vor allem für eine effiziente Preisbildung (und sind daher auch an einem PSM wichtige Marktteilnehmer), da die Effekte eines solchen Geschäfts sind, daß sich die unvereinbaren Kursdifferenzen angleichen und im

Anschluß eine weitere profitable Arbitrage mit denselben Werten nicht mehr möglich ist.

Der Faktor Zeit spielt also eine wesentliche Rolle für den Erfolg eines Arbitrageurs, der grundsätzlich von einem Spekulanten zu unterscheiden ist.

Allgemeine Definition:<sup>14</sup>

$$p(n_1 \cdot x_1 + n_2 \cdot x_2 + \dots + n_k \cdot x_k) \neq n_1 \cdot p(x_1) + n_2 \cdot p(x_2) + \dots + n_k \cdot p(x_k)$$

$p(n_1 \cdot x_1 + n_2 \cdot x_2 + \dots + n_k \cdot x_k)$ .....Preis des gehaltenen Portfolios

$x_i$ .....asset i

$n_i$ .....Anzahl der gehaltenen shares von asset i

Das heißt, die Bestandteile eines Aktienportfolios einzeln betrachtet, haben zusammen einen anderen Wert als ihr Gesamtportfolio.

Die Ähnlichkeiten dieser Definition mit dem Arbitragebegriff eines PSM sind unverkennbar:<sup>15</sup>

$$p(n \cdot x_1 + n \cdot x_2 + \dots + n \cdot x_k) = \text{Bündelpreis} \cdot n \neq n_1 \cdot p(x_1) + n_2 \cdot p(x_2) + \dots + n_k \cdot p(x_k)$$

Das heißt, ein auf Basis der Marktpreise aus Einzelaktien zusammengestelltes Basisbündel hat am Sekundärmarkt nicht den vordefinierten Wert (Bündelpreis), um den solche basic portfolios am Primärmarkt gehandelt werden.

Händler können risikolose Arbitragegewinne am APSM durch zwei mögliche Konstellationen lukrieren:

- 1) Summe der Verkaufsangebote aller Aktien (entspricht einem Bündel) eines Marktes beträgt weniger als der Bündelpreis. Dies ist der Fall wenn ein Händler ein neues bestes Verkaufsangebot abgibt und folgende Ungleichung erfüllt ist:

$$\sum A_i^t < \text{Bündelpreis}$$

$A_i^t$ .....Das beste Verkaufsangebot der Aktie i zur Zeit t

---

<sup>14</sup> nach [Jarrow 1988] S. 21ff in [Ortner 1996] S. 13

<sup>15</sup> vgl. [Ortner 1996] S. 13

Registriert ein Händler diese Situation, kann er jede einzelne Aktie am Sekundärmarkt kaufen, um anschließend dieses Bündel zum Bündelpreis mit Gewinn sofort wieder am Primärmarkt an den Marktmanager zu verkaufen. Dieser Vorgang kann so lange wiederholt werden, bis die Ungleichung nicht mehr erfüllt ist.

2) Summe der Kaufangebote aller Aktien (entspricht einem Bündel) eines Marktes ist größer als der Bündelpreis. Dies ist dann der Fall, wenn ein Händler ein neues bestes Kaufangebot abgibt und folgende Ungleichung erfüllt ist:

$$\sum B_i^t > \text{Bündelpreis}$$

$B_i^t$ ....Das beste Kaufangebot der Aktie  $i$  zur Zeit  $t$

Es gilt das gleiche wie bei 1), nur daß natürlich diesmal der Arbitrageure Bündel am Primärmarkt kauft und jede Aktie einzeln am Sekundärmarkt sofort wieder verkauft. Dies wiederholt er wieder solange bis die Ungleichung nicht mehr erfüllt ist.

### ***II.3.4 Marginal Traders***

Forsythe erkannte in seinem ersten Experiment 1988 eine kleine Gruppe von Händlern, die sich durch eine überdurchschnittliche Aktivität auf den jeweiligen Märkten auszeichnet.

Sie beobachteten den Markt genau, plazierten eine überdurchschnittliche Anzahl von Orders und nutzten schon kleine Kursschwankungen für Transaktionen. Ihre Kaufs- und Verkauforders plazierten sie sehr nahe am Marktpreis, um durch intensives Handeln Handelsgewinne zu lukrieren.<sup>16</sup>

Marginal Traders sind wesentlich an der Kursbildung und am Zustandekommen großer Umsatzvolumina verantwortlich, da sie schon bei geringen Angebots- und Nachfrageschwankungen reagieren. Obwohl eine Korrelation von Marginal Traders

---

<sup>16</sup> Ständig wird versucht, günstig Aktien aufzukaufen, um sie wenig später mit teilweise geringen Renditen wieder zu veräußern

mit hoher Rendite in späteren PSMs (z.B. im APSM95/96<sup>17</sup>) nicht immer beobachtet werden konnte, so profitieren Märkte, in denen viele Marginal Traders aktiv sind, durch:<sup>18</sup>

- ein schnelleres Ingangkommen des Handels / der ersten Kursbildung
- höheren Umsatz
- kleinere Differenz zwischen höchstem Ankaufsgebot und niedrigstem Verkaufsangebot
- schnelles Erkennen und Beseitigen von risikolosen Arbitragen
- Erkennen und Reagieren auf Manipulationsversuche
- schnelleres Antizipieren von neuen Informationen

### **//.3.5 News-Days**

Darunter versteht man wichtige Ereignisse, wie z.B.: Äußerungen eines Politikers, Vorfälle,.. die den Verlauf der Wahlschlacht maßgeblich beeinflussen. Wie stark ein solcher Einfluß ist, wird im PSM aus dem Kursverlauf ersichtlich. Dies eignet sich natürlich vorzüglich für ex-post-Analysen (vorausgesetzt natürlich, daß der Markt erfolgreich war, und die Ereignisse von den Teilnehmern „richtig“ aufgefaßt wurden).

Eine wichtige Erkenntnis im Zusammenhang mit „News“ zog Forsythe schon im IPSM88, und zwar, daß veröffentlichte Umfrageergebnisse keinen oder nur unwesentlichen Einfluß auf die Kurse im PSM haben und daher keine News darstellen. Dies wurde auch in den APSM-Projekten festgestellt.<sup>19</sup> Die vertretene Meinung einiger Sozialforscher („*Die Leute im PSM handeln ja ohnehin nur in Folge der veröffentlichten Meinungsumfragen*“<sup>20</sup>) kann somit nicht bestätigt werden.

---

<sup>17</sup> vgl. [Pawek 1996] S. 32

<sup>18</sup> vgl. [Ortner 1996] S. 56

<sup>19</sup> vgl. [Ortner 1996] S. 42,43

<sup>20</sup> Georg Michenthaler von IFES im Interview vom 28.10.1997

## **II.4 AEM/APSM97**

Da die Abteilung Industrielle Betriebswirtschaftslehre 1997 den elektronischen Markt auch für andere Prognosen verwendete, wie etwa den Markt „ALR797 – Arbeitslosenrate Juli 1997“, nannte man die Internetseite nicht mehr APSM sondern AEM (Austrian Electronic Markets). Der Markt „Oberösterreichische Landtagswahl 1997“ lief also unter AEM/APSM97. Im folgenden wird dieses Projekt allerdings nur mehr APSM97 genannt.

### **II.4.1 Design des APSM97**

Wie schon bei den Projekten vergangener Jahre wurde der Markt ähnlich des Fließhandels an der Wiener Börse als „continuous double auction market“ aufgesetzt. Das heißt, ein Handel war rund um die Uhr an sieben Tagen in der Woche bis zum Schließen der Märkte am Wahltag (5. Oktober 97) möglich. Die Teilnahme stand allen interessierten österreichischen Internetusern offen.

Voraussetzung für die Teilnahme war das Ausfüllen und Versenden der Anmeldung an das Marktmanagement, das Einzahlen eines Investitionsvolumens (100,- bis 5000,- öS) und das Einverständnis mit den Modalitäten des Experiments, welche waren:

- Das Einverständnis, daß die nicht personenbezogenen bzw. anonymisierten (wie z.B. die abgeschlossenen Handel, die Volumina,..) zu wissenschaftlichen Auswertungen verwendet werden.
- Die Auszahlungsregeln – siehe später.
- Der Charakter eines Experiments
- Die Rückzahlung des eingesetzten Kapitals hing ausschließlich vom Geschick des Teilnehmers ab.

- Es wurde keine Haftung für eventuell nicht reibungslos verlaufene Handel wegen Hard- und Softwarefehler übernommen.

Nach der Anmeldeerklärung und der Einzahlung des Investitionsbetrags (Eingang des Betrages auf dem Konto) erhielt ein Teilnehmer ein Computerkonto (UserID, Paßwort), daß ihm ermöglichte, in den Märkten zu handeln. UserID und Paßwort wurden per Email den Teilnehmern mitgeteilt. Damit verfügte ein Teilnehmer über ein eigenes Aktiendepot, auf dem zunächst nur der eingezahlte Geldbetrag (Cashposition) deponiert war.

Die Marktteilnehmer konnten frei entscheiden, auf welchen Märkten sie ihr Geld investieren wollten. Zum Abgeben eines Kauf- bzw. Verkaufsorders waren folgende Angaben notwendig: Welcher Markt (ÖVP, SPÖ,..), Limits für den Verkauf- oder Kaufpreis und die Anzahl. Noch nicht (komplett) ausgeführte Aufträge konnten jederzeit zurückgezogen (deleted) werden und wurden dann auch nicht ausgeführt.

#### Auszahlungsregeln – Vote Share Market:

Die Auszahlung erfolgte in Abhängigkeit vom Ausgang der Wahlen. Nach der geschlagenen Wahl wurden alle Aktien des betreffenden Marktes vom Marktmanagement eingezogen/zurückgekauft. Dabei wird für jede Aktie ein Preis vom Stimmenanteil der entsprechenden Partei \* 1 öS vergütet (Ein basic portfolio im APSP97 betrug 1öS). Das heißt, für ein komplettes Aktienbündel (basic portfolio) werden in Summe wieder genau 1 öS bezahlt, da ein solches Bündel ja genau 100% der Stimmen widerspiegelt.

Besitzt ein Händler z.B. am Stichtag 2 X- und 4 Y-Aktien und das Ergebnis der fiktiven Wahl zwischen den Parteien X und Y möge 55 zu 45% sein, dann beträgt die Auszahlung:  $2 * (0,55 * 1\text{öS}) + 4 * (0,45 * 1\text{öS}) = 2,9\text{öS}$



Software:

Der APISM97 lief auf einem WWW-Server der Abteilung „Industrielle Betriebswirtschaftslehre“. Die verwendete Software wurde in Canada von der Firma Kumo Inc. entwickelt und 1995 bei der Ars Electronica ausgezeichnet.

Bei früheren Experimenten wurde eine Software verwendet, die auch bei den Experimenten in Iowa zur Anwendung gelangt. Probleme gab es dabei aber vor allem mit der Abwicklungsgeschwindigkeit der Transaktionen. Es konnte passieren, daß ein Marktteilnehmer so lange warten mußte bis seine Befehle ausgeführt wurden, daß er die eventuell gewollten Transaktionen abbrach.

**Durch die Verwendung der neueren, kanadischen Software gab es aber auch einige Veränderungen im Marktdesign:**

1) Handel in fünf verschiedenen Märkten:

Während es bei den APISM94-96 immer einen Markt gab, in dem dann alle Aktien (SPÖ, ÖVP, FPÖ, Grüne, LiF und Andere) gehandelt werden konnten, gab es im APISM97 fünf verschiedene Märkte, für jede im Nationalrat vertretene Partei einen.

Die Andere-Aktie konnte nicht direkt gehandelt werden, dieser Wert ergab sich also aus 100% weniger der Summe der Kurse der fünf Parteien.

Durch die Aufspaltung in fünf Märkten ist die Übersichtlichkeit natürlich stark eingeschränkt. Die Bids und Asks für jede Aktie müssen einzeln abgeprüft werden, das Gefühl für den Gesamtmarkt nimmt ab, Arbitragemöglichkeiten sind nicht leicht erkennbar, sondern erfordern einen gewissen Aufwand.

Positiv ist zu bemerken, daß bei Einstieg in einen Markt, interessante Informationen wie Kursverlauf in graphischer Form zur Verfügung standen. Dies verbesserte den Informationsstand jedes Marktteilnehmers.

## 2) Bündelpreis:

Wie schon erwähnt betrug der Preis für ein komplettes Aktienbündel (basic portfolio) 1 öS, bei früheren Experimenten 10 öS. Dadurch wird in bestimmter Weise einer Annahme des vollkommenen Kapitalmarkts Rechnung getragen, und zwar, daß Rechte an Wertpapieren beliebig teilbar sind. Es ist dadurch möglich, schon mit sehr kleinen Investitionen in einen Markt einzusteigen.

Der alte Bündelpreis betrug zwar 10öS, es war aber möglich, einzelne Aktien auf bis zu drei Stellen hinter dem Komma zu handeln, z.B. eine LiF-Aktie um 0,361 – entspricht einem Stimmenanteil von 3,61%.

Im APSM97 waren die Aktien nur auf ganze Prozent zu handeln, z.B. eine LiF-Aktie um 3 oder 4 Groschen – entspricht einem Stimmenanteil von 3 oder 4%. Dies ist natürlich besonders bei kleinen Parteien äußerst negativ zu bewerten, denn der Sprung von 3 auf 4% beträgt 33%. Der Informationsstand über den Preis einer Aktie – z.B. eben 3,61% - konnte also nicht optimal verwertet werden, da man dazu die Aktie entweder um 3% kaufen oder um 4% verkaufen müßte.

## 3) Keine eindeutige Erkennbarkeit des Primärmarkts

Während bei früheren Projekten die Teilnehmer zuerst basic portfolios am Primärmarkt kaufen mußten um einzelne Aktien am Sekundärmarkt zu verkaufen, liefen diese beiden Vorgänge beim APSM97 „automatisch“ durch den Handel am Sekundärmarkt ab, ohne daß dem Teilnehmer dies direkt verdeutlicht wurde.

Technisch wurde dies dadurch ermöglicht, daß ein Teilnehmer immer ein Paar (Ja, Nein) von Anteilen kaufte, wenn er am Sekundärmarkt Transaktionen durchführen wollte.

Verkaufte ein Teilnehmer z.B. 10 Stück einer Aktie, die er nicht besaß, um je 40 Groschen am Sekundärmarkt (er glaubte also, daß die Partei bei der Wahl unter 40% erzielen werde), liefen automatisch folgende Vorgänge ab:

- Zehn Paare (Ja, Nein) von Anteilen, dessen Preis ja immer 100% beträgt, wurden am Primärmarkt gekauft. Die Cash balance verringerte sich damit um 10 öS. (=  $10 \cdot \text{Bündelpreis}$ )
- Dann wurden die Ja-Anteile sofort wieder um 40 Groschen verkauft. Die Cash balance stieg damit wieder um 4öS.
- Der Teilnehmer behielt die Nein-Anteile. Nehmen wir an, die Aktie war nach dem Schließen des Marktes 38 Groschen wert (Ausgang der Wahl: 38% für diese Partei). Dann wurden  $100 - 38 = 62$  Groschen pro Nein-Aktie gutgeschrieben. Dadurch erhöhte sich der Cash balance um 6,2öS. Erfolg der Transaktion: 0,2öS

Dies erleichterte den Handel am Sekundärmarkt natürlich wesentlich, und ist als signifikante Verbesserung des Design einzustufen.

#### 4) Zeitlich limitierte Order

Bei der in Iowa entwickelten Software konnte man Orders zeitlich limitieren, das heißt daß diese dann automatisch zum gewählten Zeitpunkt vom Markt zurückgezogen wurden, wenn sie bis dahin nicht ein anderer Marktteilnehmern annahm.

Diese zeitliche Limitierung war im APSM97 nicht mehr möglich. Besonders Marginal Traders haben in früheren Projekten diese zeitliche Limitierung stark genutzt. Ihnen ist dieses Instrument also „abhanden gekommen“.

#### 5) Offenes Orderbook

1997 wurde die open order-book Strategie verfolgt. Damit meint man eine Auflistung aller aktuellen Bids und Asks, die es im jeweiligen Markt gerade gibt. Wenn man also in einen Markt einsteigt, könnte man die rechten Äste der Angebots- und der Nachfragekurve aufzeichnen (die linken Teile der beiden Kurven fallen ja erst gar

nicht an, weil bei einer Überschneidung von Bids und Asks sofort eine Transaktion folgt). In den früheren Projekten war dies nicht möglich, da nur das jeweils beste Angebot ersichtlich war.

Dadurch wird der Informationsstand jedes Traders deutlich verbessert. Die zentrale Frage ist aber, ob dies dem Markt und dem schnellen Einstellen von Marktpreisen dienlich ist. Der wesentlichste Voraussetzung in diesem Zusammenhang ist, daß die Markttiefe, also die Liquidität des Marktes ausreichend vorhanden ist.

Gibt es in einem Markt wenig Trader, und neu einsteigende können dies natürlich über das offene Orderbook sofort erkennen, dann sinkt der Anreiz Transaktionen durchzuführen, es werden eher zuerst Orders abgegeben, die nicht sofort zu Transaktionen führen.

*„Of course, if there are a small number of liquidity suppliers, then there is an incentive for others to provide quotes.“<sup>21</sup>*

Auf der anderen Seite wird gerade dadurch die Liquidität eines Marktes erhöht, da auch risikoaverse Teilnehmer einmal eine Order plazieren, weil die „Gefahr“, ob dies sofort zu einer Transaktion führt, abschätzbar ist.

In realen Märkten, wie z.B. der London Stock Exchange (Ein Händlermarkt) oder der Pariser Börse (elektronisches offenes Orderbook) kann eine Tendenz beobachtet werden, daß Teilnehmer, die in beiden Märkten aktiv sind, eher zum Händlermarkt tendieren. Dies kann aber natürlich auch an der Attraktivität des London Stock Exchange liegen. Für den APSM97 kommt diese Diskussion natürlich nicht so in Betracht, da man keinen konkurrierenden PSM hatte, es sei den manche Teilnehmer handeln auch am IPSM oder an einem anderen PSM ohne offenen Orderbook. Dies wird aber auf die meisten Teilnehmer nicht zutreffen.

---

<sup>21</sup> vgl. [Glostein 1994] S. 1152

Zusammenfassend kann gesagt werden, daß das Hauptproblem des offenen Orderbooks das langsame „Anlaufen“ des Marktes ist. Einen positiven Effekt würde es also nur haben, wenn großes Interesse am Markt besteht, die Markttiefe ausreichend vorhanden ist, und vom Marktmanagement eine erste Preisbildung initiiert wird.

Da noch zuwenig praktische Erfahrung mit der Strategie des offenen Orderbooks in PSMs vorliegt, müssen die zukünftigen Ergebnisse kritisch betrachtet werden, um eine klare Aussage über die Effekte auf das Marktdesigns machen zu können.

#### ***II.4.2 Teilnehmeranalyse***

Das Teilnehmerfeld war im APSM97 in Größe anderen Projekten deutlich unterlegen. Es gab nur 10(!) aktive Trader, das allgemeine und mediale Interesse an der Oberösterreichischen Landtagswahl war nicht gerade groß, schon gar nicht außerhalb Oberösterreichs, das Experiment lief aber an der TU Wien.

Unter den neuen Marktteilnehmern wurde eine anonyme Meinungsumfrage durchgeführt. Es gab aber auch einige Teilnehmer, die auch schon bei früheren Projekten teilnahmen, und in den APSM97 „wechselten“. Von diesen gab es natürlich keine Umfrageergebnisse. Die Zahl der neuen Teilnehmer, von denen Umfrageergebnisse existieren betrug zehn, also zufällig genau die Zahl der aktiven Trader. Obwohl nicht gesichert ist, daß genau diese zehn auch die zehn aktiven waren, wird die Meinungsumfrage nun analysiert, es sind aber natürlich Verzerrungen zum tatsächlichen, unbekanntem aktiven Teilnehmerkreis möglich.

Die Abbildungen 2.9-2.17 zeigen die Ergebnisse der Meinungsumfrage unter den neuen Marktteilnehmern.

# Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Abbildung 2.9: Wie oft in der Woche lesen Sie den politischen Teil einer Tageszeitung?

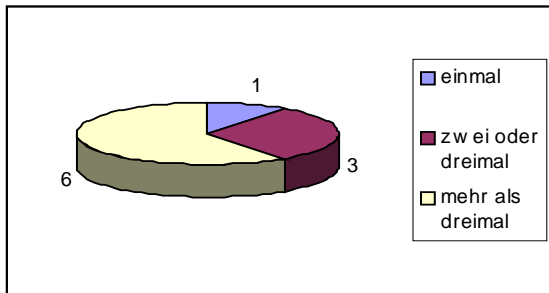


Abbildung 2.10: Wie oft in der Woche lesen Sie den Wirtschaftsteil einer Tageszeitung?

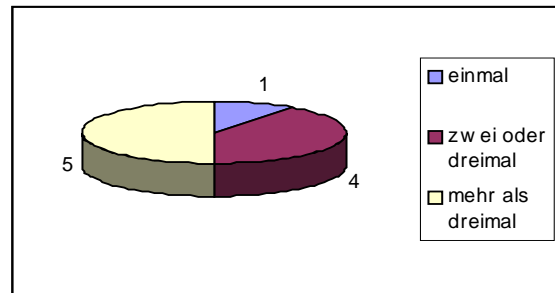


Abbildung 2.11: Wie oft in der Woche verfolgen Sie durchschn. die Nachrichten im Fernsehen?

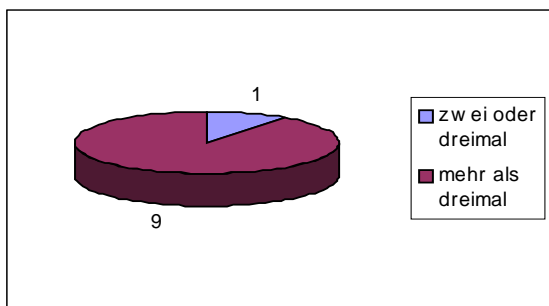


Abbildung 2.12: Wenn heute Nationalratswahlen wären, welche Partei würden Sie wählen?

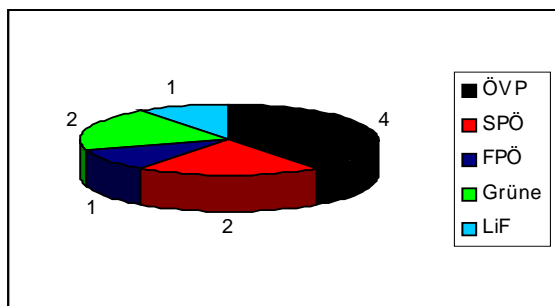


Abbildung 2.13: Sind Sie am „normalen“ Börsengeschehen interessiert?

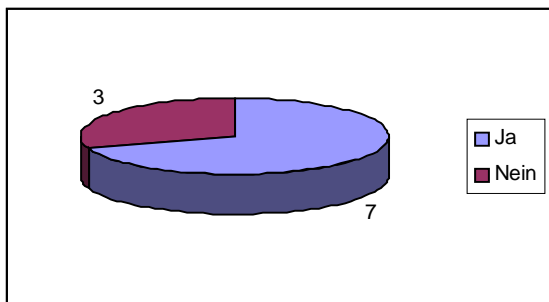


Abbildung 2.14: Zur Person

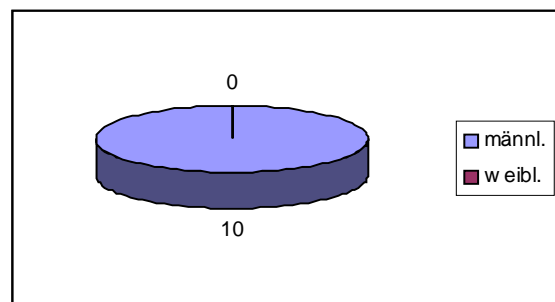


Abbildung 2.15: Beruf

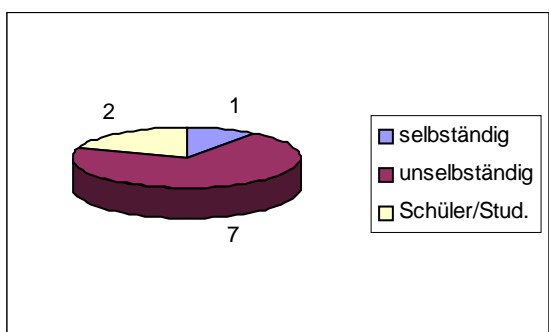


Abbildung 2.16: Alter

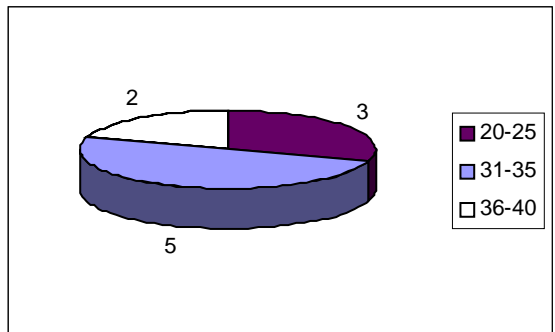
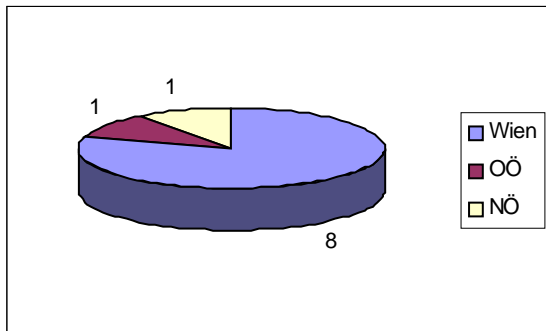


Abbildung 2.17: In welchem Bundesland Wohnen bzw. leben Sie?



Es fällt natürlich das große Interesse der Teilnehmer am politischen und wirtschaftlichen Geschehen auf, wie man aus den Ergebnissen der Fragen 1,2,3 sehen kann. Das Interesse der überwiegenden Zahl der Teilnehmer am „normalen“ Börsengeschehen ist auch relativ hoch.

Die Repräsentativität der Teilnehmer ist in einem PSM, wie schon erwähnt, nicht erforderlich, beeinflusst aber trotzdem die Güte des Ergebnisses. Vor allem, daß nur ein einziger Teilnehmer aus Oberösterreich war, ist natürlich sehr negativ zu betrachten, denn woher beziehen Wiener, die den überwiegenden Anteil der Teilnehmer darstellten, ihre Informationen über den Ausgang der Wahl.

Weiters waren alle Teilnehmer ausschließlich Männer. Auch hier stellt sich die Frage, inwieweit Männer über das Wahlverhalten der Frauen, das ja meistens unterschiedlich ist<sup>22</sup>, informiert sind.

Es kann also schon nach dieser Teilnehmeranalyse gesagt werden, daß es nicht verwundert, wenn das Ergebnis des APSM nicht überzeugend ist, da die Anzahl und die Zusammensetzung der Teilnehmer die Qualität wesentlich beeinflussen.

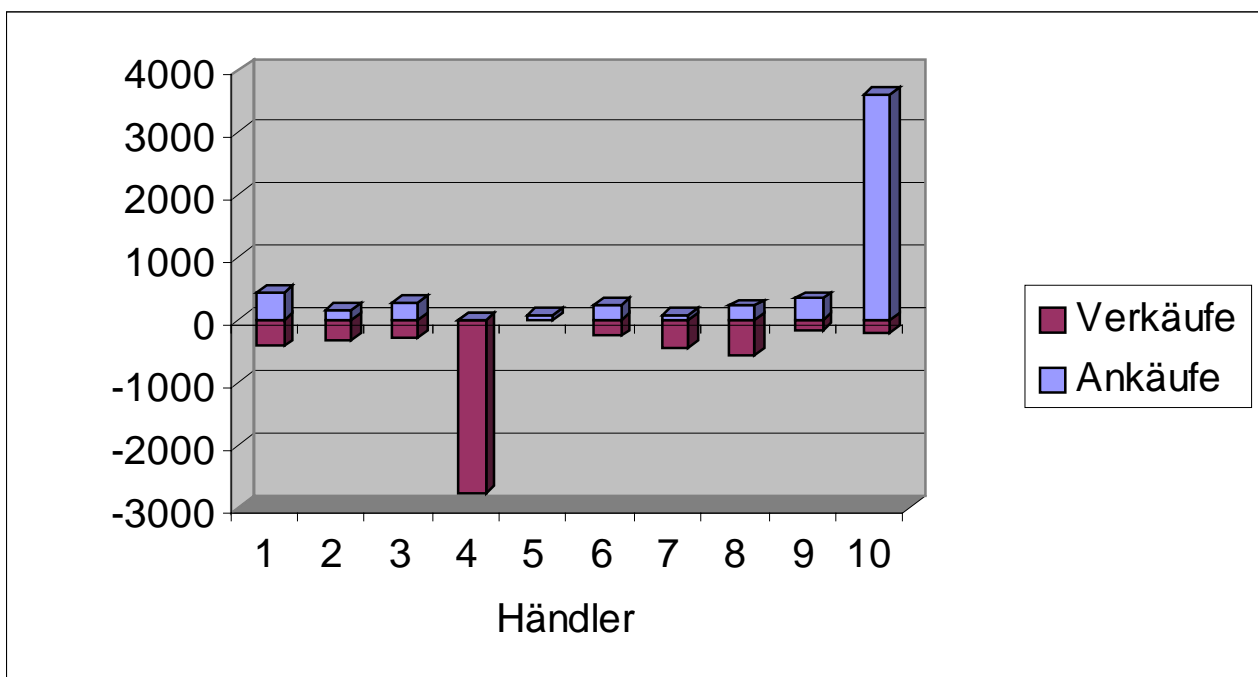
Da die Umfrage anonym durchgeführt wurde, können verschiedene Einflüsse, wie z.B. Parteipräferenz und Wahlverhalten, nicht analysiert werden. Im Vergleich dazu wurde bei den amerikanischen Marktexperimenten oft laufend Umfragen unter den

Teilnehmern durchgeführt, wobei die Auswertungen jeder Person zuordenbar waren. Dabei konnten traderspezifische Verhalten natürlich ohne weiteres untersucht werden. Dies ist bei dieser Teilnehmeranalyse leider nicht möglich.

### II.4.3 Händlerverhalten

Bei der folgenden Analyse werden den Händlern einfach die Nummern eins bis zehn zugeordnet, die Zuordnung sagt nichts über die Händler-ID oder sonstige Merkmale aus.

Abbildung 2.18: Gehandelte Aktien (Käufe und Verkäufe) in Stück



Die Dominanz der Trader 4 und 10 fällt sofort auf. Ihre Transaktionen machen 60% der gesamten aus. Während Trader 10 vorwiegend Aktien kaufte, hat Trader 4 überhaupt nur Verkäufe durchgeführt.

---

<sup>22</sup> vgl. auch Kapitel III



Die restlichen Händler haben sich eher zurückgehalten, auch ist bei ihnen keine so eindeutige Präferenz für Käufe bzw. Verkäufe zu erkennen.

Abbildung 2.19: Gehandelte Aktien (Käufe) in Stück

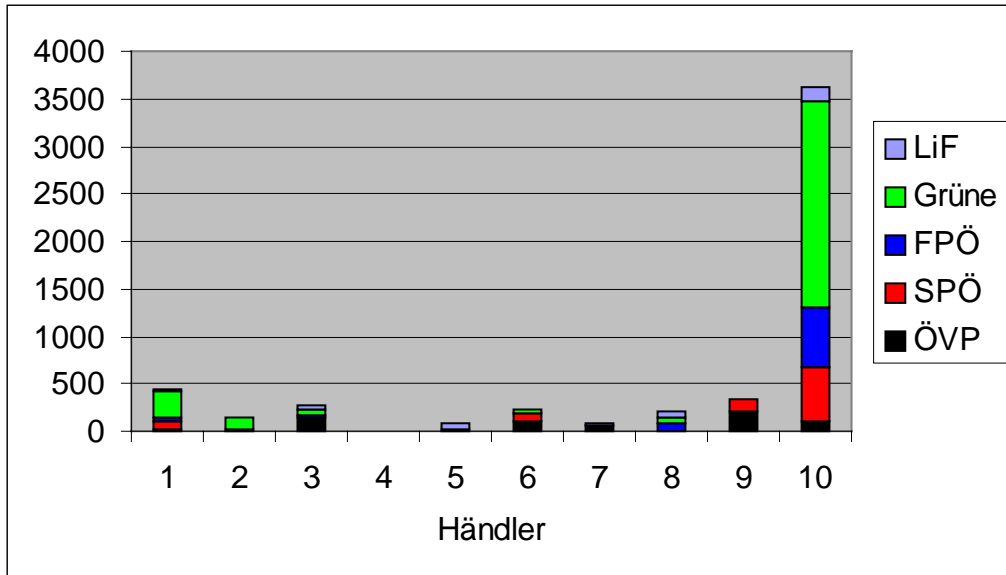
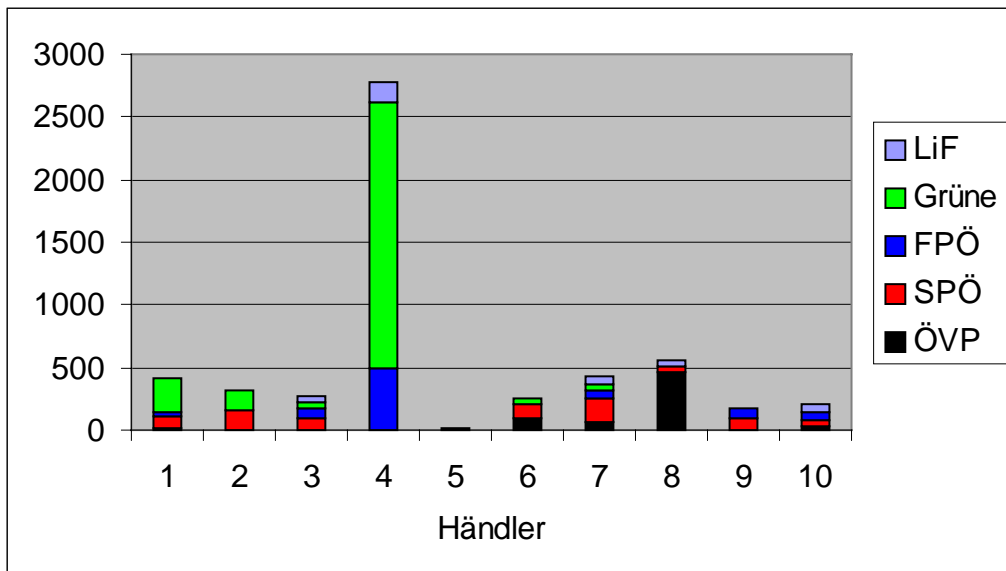


Abbildung 2.20: Gehandelte Aktien (Verkäufe) in Stück

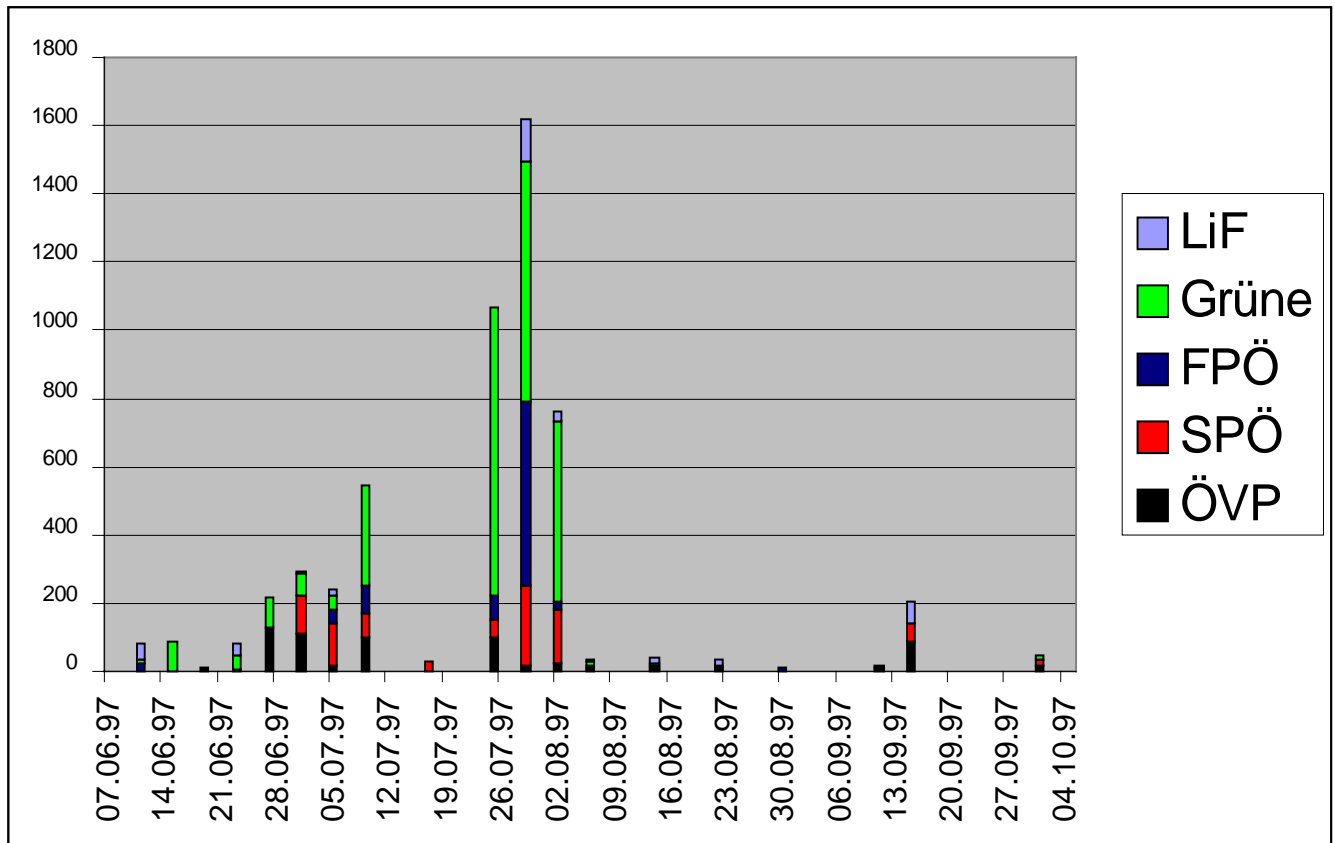


Die „Grüne“-Aktie war, wie aus Abbildung 2.18 und 2.19 ersichtlich, die beliebteste, vor allem Trader 4 und 10 haben hauptsächlich mit dieser gehandelt. Auffällig ist

## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

weitere, daß Trader 8 vor allem die ÖVP-Aktie verkauft hat und zwar zu einem eher schlechten Preis, er hat also zur Unterbewertung der ÖVP-Aktie beigetragen.

Abbildung 2.21: Gehandelte Aktien in Stück und Vier-Tages-Intervall



Von den insgesamt 5440 gehandelten Stück wurden 3448, also über 63% in der Zeit vom 22.07 bis 01.08 gehandelt. Die „Grüne“-Aktie war mit 2080 Stück in dieser Zeit die mit Abstand am meisten gehandelte, gefolgt von der „FPÖ“-Aktie. Trader 10 hat alle Grüne-Aktien vom 22.07. bis 01.08. gekauft, Trader 4 die meisten verkauft.

Trader 10 hat aber auch alle anderen Aktien, mit Ausnahme der ÖVP-Aktie, in dieser Zeit gekauft, da die Märkte aber überbewertet waren, Arbitrage war durch Verkäufe von gesamten Bündel möglich<sup>23</sup>, ist diese ökonomisch irrationale Handlung schwer nachvollziehbar. Dieser Teilnehmer hat natürlich die Preisbildung stark beeinflusst,

<sup>23</sup> vgl. auch Kapitel II.3.3 und II.4.4

und zu einer tendenziellen hohen Bewertung der Aktien SPÖ, FPÖ, Grüne und LiF beigetragen.

Alle anderen, vor allem aber Trader 4, konnten davon profitieren.

Die Motive, besonders warum gerade im Ende Juli, Anfang August so intensiv gehandelt wurde, seien dahingestellt. Besonders am Ende des Marktes flaute das Interesse im Gegensatz zu den früheren Projekten, in denen es meist noch ein deutliches Ansteigen des Handelsvolumen gab, deutlich ab. Da der Informationsstand der Marktteilnehmer am Ende eines Marktes aber deutlich größer sein müßte wie am Beginn, der Wahlausgang steht ja kurz bevor, ist dies ein negativer Indikator für das Ergebnis des APSM97.

Zusammenfassend kann gesagt werden, daß das äußerst bescheidene Handelsvolumen (5.440 gehandelte Stück im Vergleich zu 65.537 im Nationalratwahlmarkt 95) und vor allem die Dominanz von zwei Trader sehr negativ zu beurteilen ist.

Wie stark die Dominanz in den einzelnen Märkten und ihre Auswirkungen werden für jede Aktie später getrennt behandelt.

## **II.4.4 Ergebnisse und Auswertungen der Märkte**

### a) Allgemeine Resultate

In Tabelle 2.7 werden neben dem Total Error auch der MAE und RSME berechnet. Unter MAE versteht man den Mean Average Error,  $MAE = TE/n$  und unter RSME den Root Square Mean Error,  $RSME = \sqrt{((\sum d_i^2)/n)}$ .

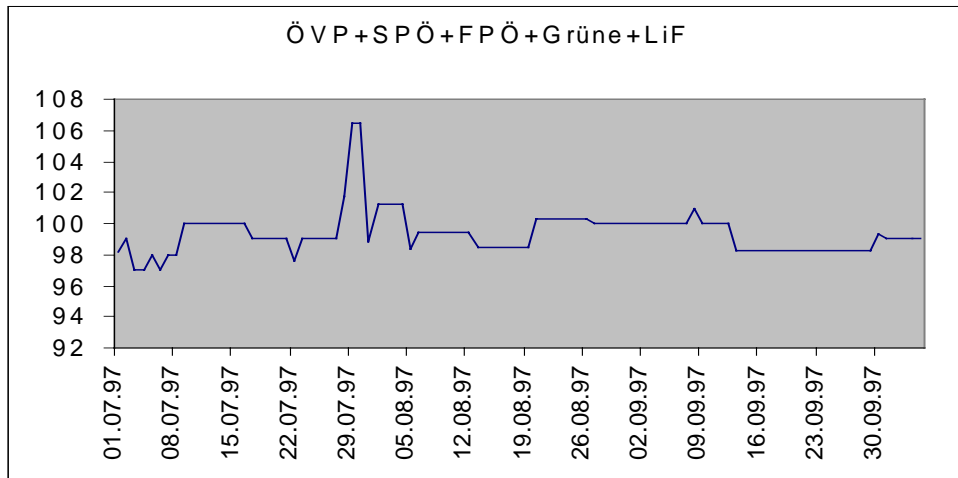
n gibt die Anzahl der verschiedenen Aktientitel, also der Parteien an, in diesem Falle ist  $n=5$ . Der RSME-Wert gewichtet höhere Einzelfehler stärker und ist daher ein sehr interessanter Parameter. Da die quadratischen Abweichungen auch durch die Anzahl der Aktientitel dividiert wird, ermöglicht dieser Parameter wie der MAE einen objektiven Vergleich, auch wenn bei einer der beiden Methoden die „Andere“-Aktie bzw. die „anderen Parteien“ nicht gehandelt bzw. prognostiziert wird.

Tabelle 2.7: Ergebnisse des APSM97 verglichen mit dem tatsächlichen Endergebnis

Partei	Ergebnis	APSM	d
ÖVP	42,69	40	2,69
SPÖ	27,04	31	3,96
FPÖ	20,63	22	1,37
Grüne	5,78	7	1,22
LIF	2,09	4	1,91
Andere	1,77	-	-
TE			11,15
MAE			2,23
RSME			2,45

Das Ergebnis ist äußerst bescheiden, wie schon erwähnt, in Hinblick auf die geringe Teilnehmerzahl und deren Zusammensetzung nicht weiter überraschend.

Abbildung 2.22: Zeitlicher Verlauf der Summe der Kurse aller Aktien



Wie aus Abbildung 2.21 ersichtlich, lag die Summe der Kurse stets nahe bei 100% oder sogar darüber. Da nur maximal 100% rational sind, sogar eher weniger, denn andere Parteien erkämpfen ja auch immer kleine Anteile, wird deutlich, daß mögliche Arbitragegewinne nicht ausreichend genutzt wurden. Man muß natürlich berücksichtigen, daß es zu einem Zeitpunkt möglicherweise keine Kauforders zum jeweiligen Kurs der Aktien gegeben hat.

b) ÖVP-Aktie

Abbildung 2.23: Kursverlauf und Handelsvolumen im ÖVP-Markt

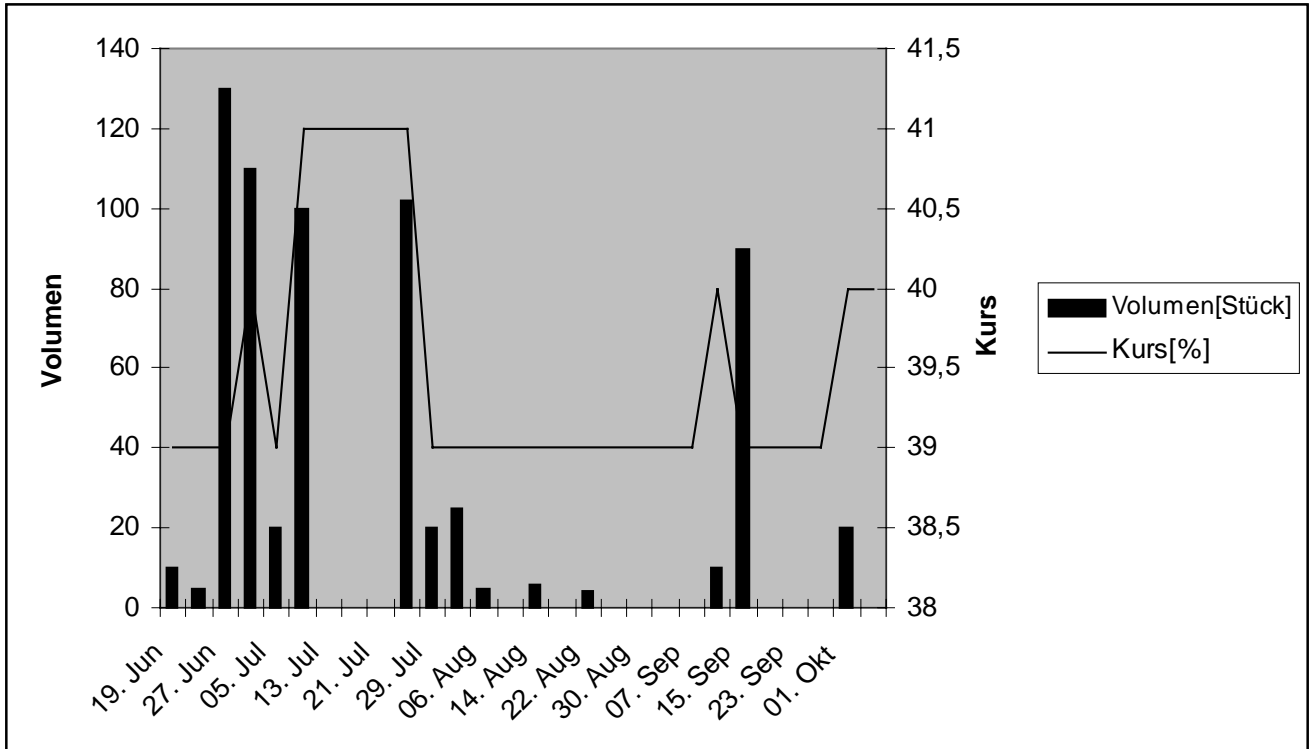
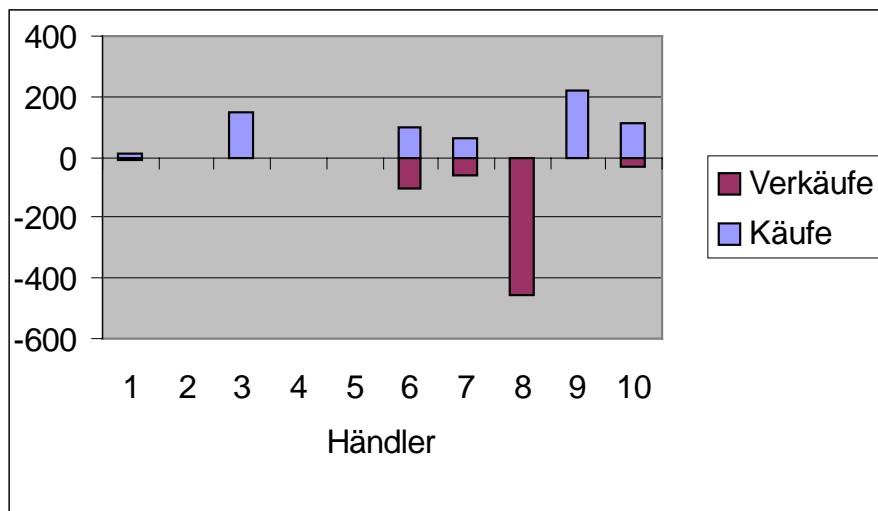


Abbildung 2.24: Gehandelte Aktien (Käufe und Verkäufe) – ÖVP



Die ÖVP-Aktie schwankte zwischen 39 und 41% und schloß mit 40% am Ende des Marktes. Damit war sie um etwas mehr als 2% unterbewertet.

Sie wurde vor allem vor der Sommerpause stark gehandelt, Ende Juli, Anfang August und dann noch einmal Mitte September. Ansonsten war das Interesse an dieser Aktie, verglichen mit den anderen Titeln, gering.

Trader 8 spielte eine dominante Rolle im ÖVP-Markt, er hat den Großteil der Käufe getätigt, und dies zu einem Preis um 39, 40. Er hat damit den Kurs maßgeblich beeinflusst, vor allem Trader 3,9 und auch 10 konnten von der Unterbewertung der ÖVP-Aktie durch Trader 8 Nutzen ziehen.

### c) SPÖ-Aktie

Abbildung 2.25: Kursverlauf und Handelsvolumen im SPÖ-Markt

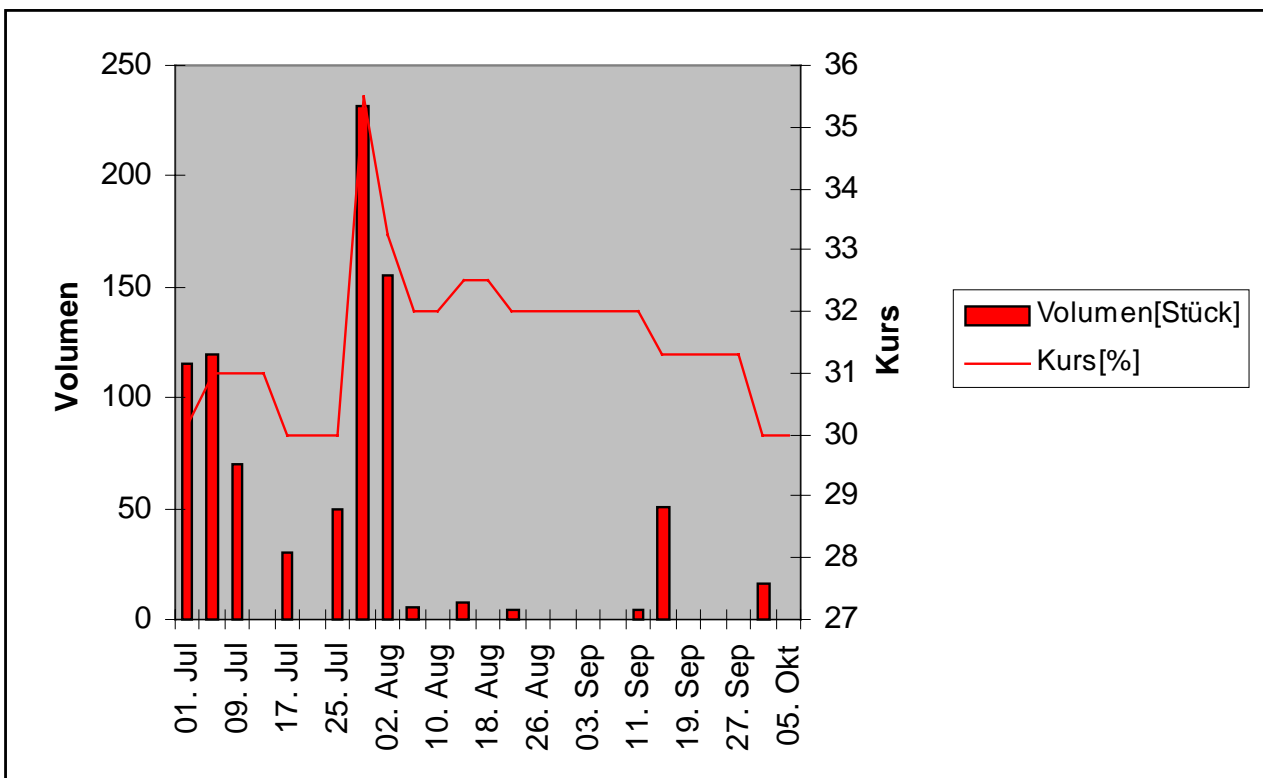
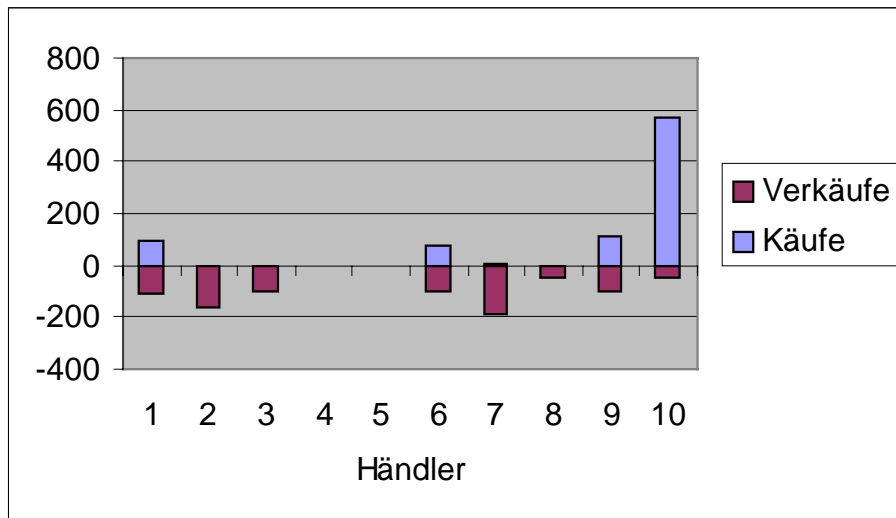


Abbildung 2.26: Gehandelte Aktien (Käufe und Verkäufe) – SPÖ



Die SPÖ-Aktie schwankte zwischen 30 und über 35% und schloß mit 30% am Schlußtag. Sie war damit um fast 3% überbewertet.

Sie wurde Anfang, Mitte Juli, Ende Juli und Anfang August stark gehandelt, das Interesse war insgesamt höher wie an der ÖVP-Aktie.

Trader 10 spielte eine dominante Rolle in diesem Markt. Er hat den Großteil der Käufe getätigt, damit zur Überbewertung der SPÖ-Aktie stark beigetragen. Vor allem Trader 2,3 und 7 konnten davon profitieren.



d) FPÖ-Aktie

Abbildung 2.27: Kursverlauf und Handelsvolumen im FPÖ-Markt

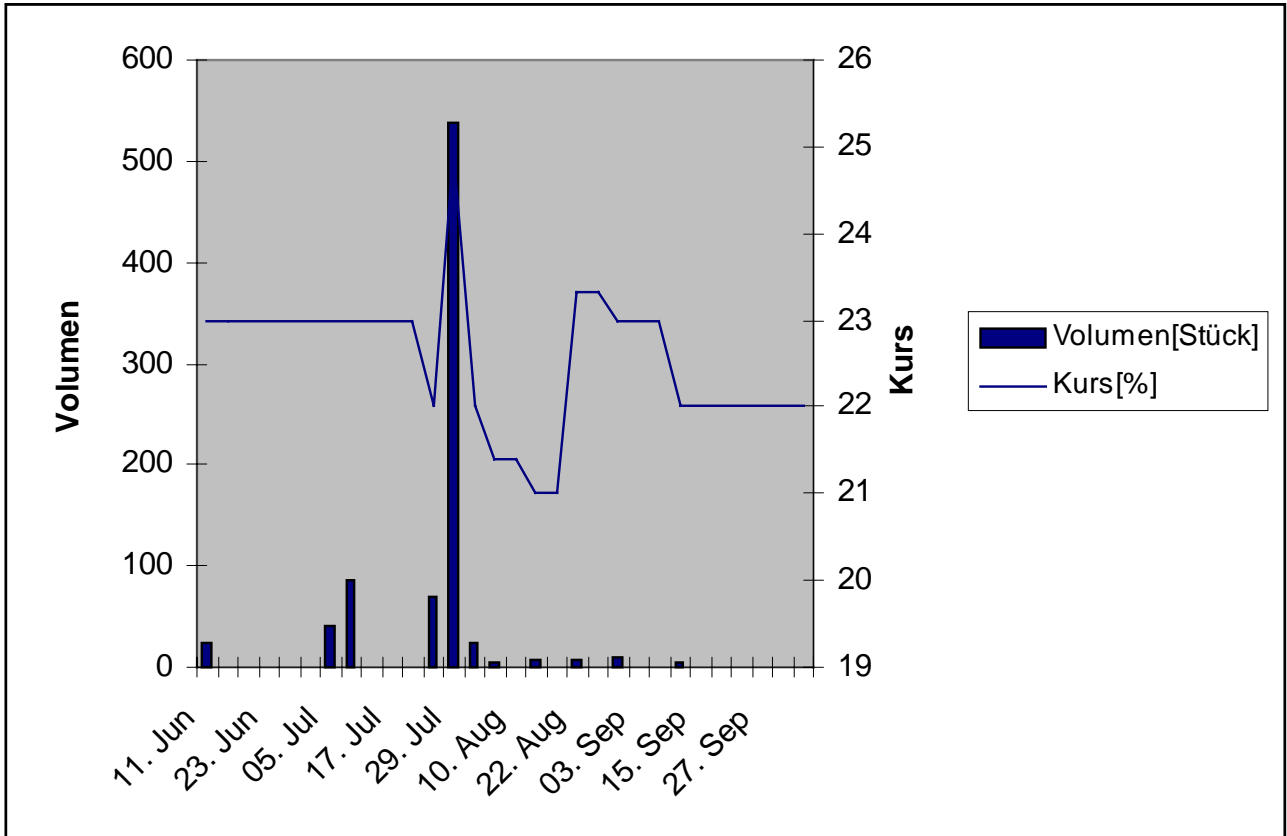
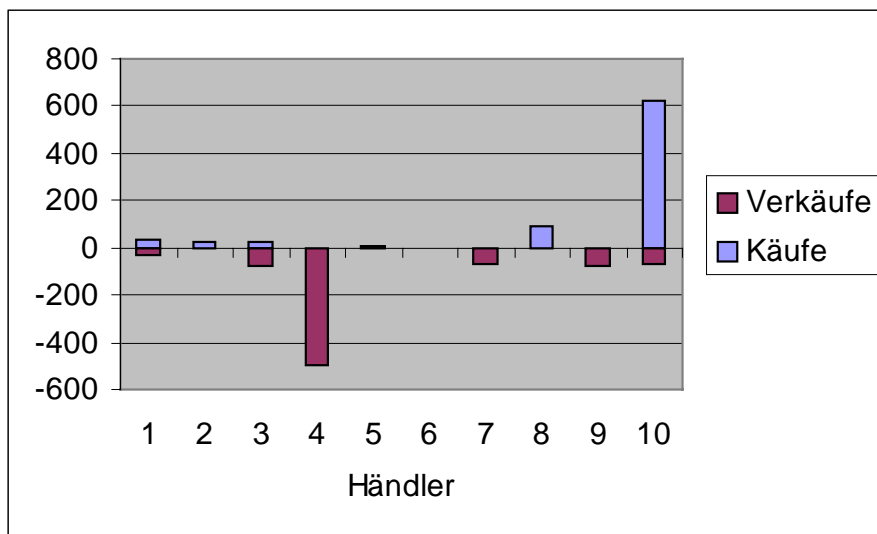


Abbildung 2.28: Gehandelte Aktien (Käufe und Verkäufe) – FPÖ



Der Kurs der FPÖ-Aktie schwankte zwischen 21,4 und fast 25% und schloß mit 22% am letzten Tag. Damit war die Aktie leicht überbewertet.

Sie wurde Anfang Juli und vor allem Ende Juli, Anfang August stark gehandelt.

Trader 4 als Verkäufer und Trader 10 als Käufer spielten eine dominante Rolle im FPÖ-Markt, wobei diese beiden hauptsächlich Ende Juli, Anfang August handelten, die Aktie wurde hier sehr hoch gehandelt und erreichte ihren höchsten Wert. Trader 4 konnte also stark von der Überbewertung durch Trader 10 profitieren.

Danach wurde die Aktie kaum mehr gehandelt, der Schlußkurs war also Ergebnis weniger Transaktionen Mitte, Ende August.

#### e) Grüne-Aktie

Abbildung 2.29: Kursverlauf und Handelsvolumen im Grüne-Markt

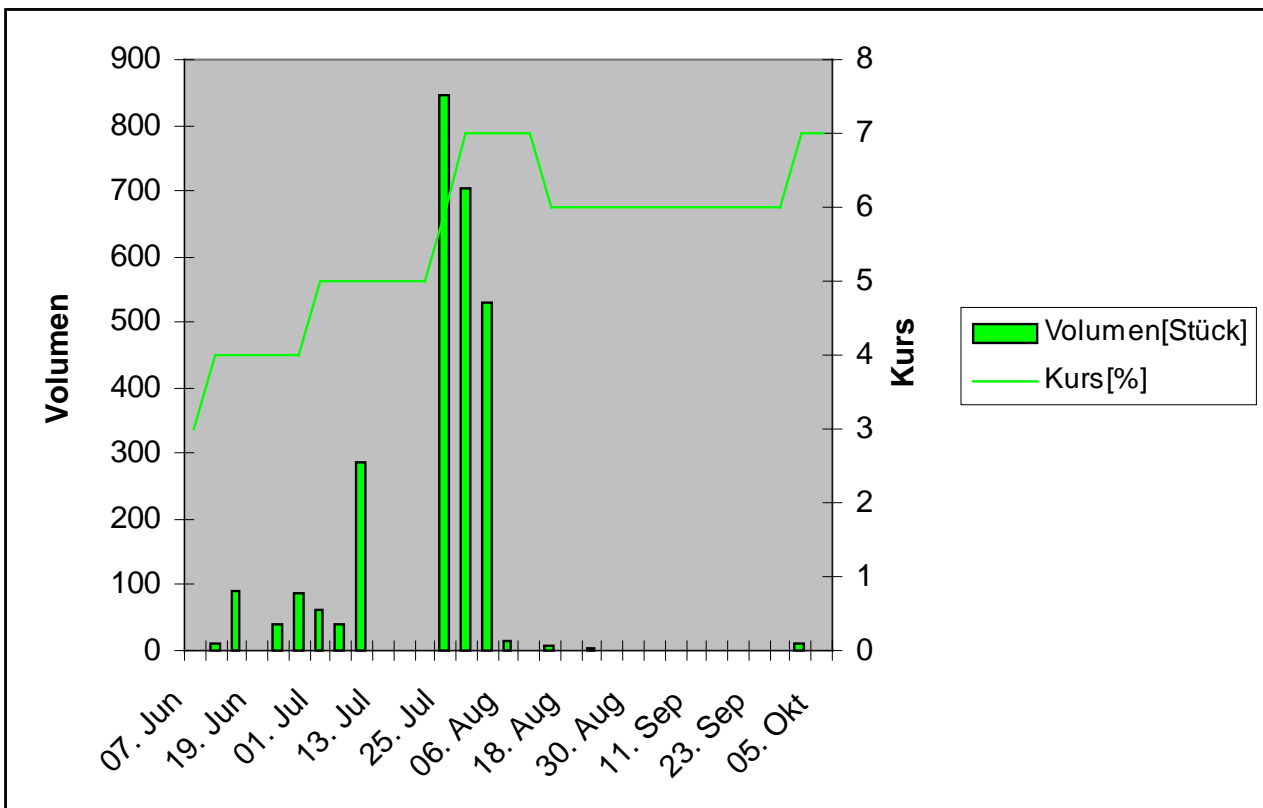
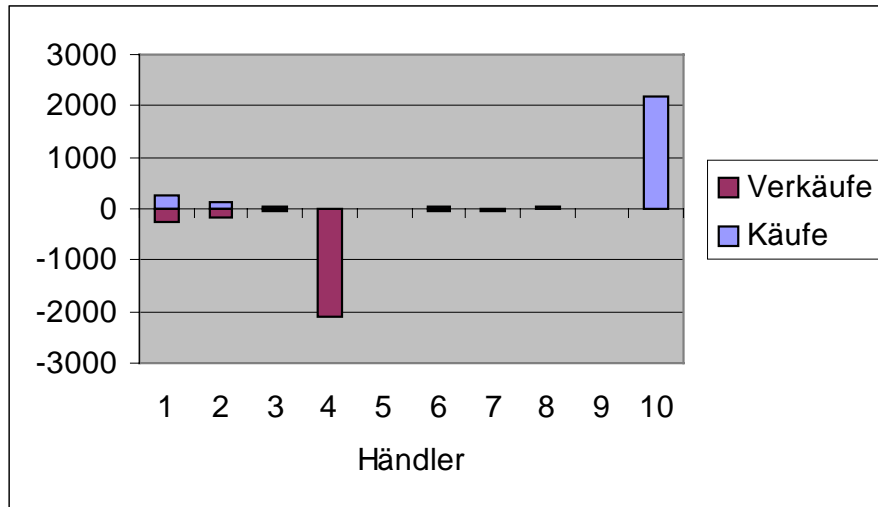


Abbildung 2.30: Gehandelte Aktien (Käufe und Verkäufe) – Grüne



Der Kurs der Grüne-Aktie schwankte zwischen 3 und 7% und schloß mit 7% am Schlußtag. Sie war damit um etwas über 1% überbewertet.

Der Großteil der Transaktionen spielte sich Mitte Juni bis Mittel Juli und auch Ende Juli, Anfang August ab. Danach wurde die Aktie kaum mehr gehandelt. Trotzdem war die Grüne-Aktie jene, die mit Abstand am meisten gehandelt wurden.

Händler 4 und 10 waren für den Hauptteil der Transaktionen verantwortlich, hauptsächlich Ende Juli, Anfang August. Sie waren sehr dominant in diesem Markt. Der Schlußkurs beruht aber aufgrund weniger Transaktionen Ende September, Anfang Oktober.

f) LiF-Aktie

Abbildung 2.31: Kursverlauf und Handelsvolumen im LiF-Markt

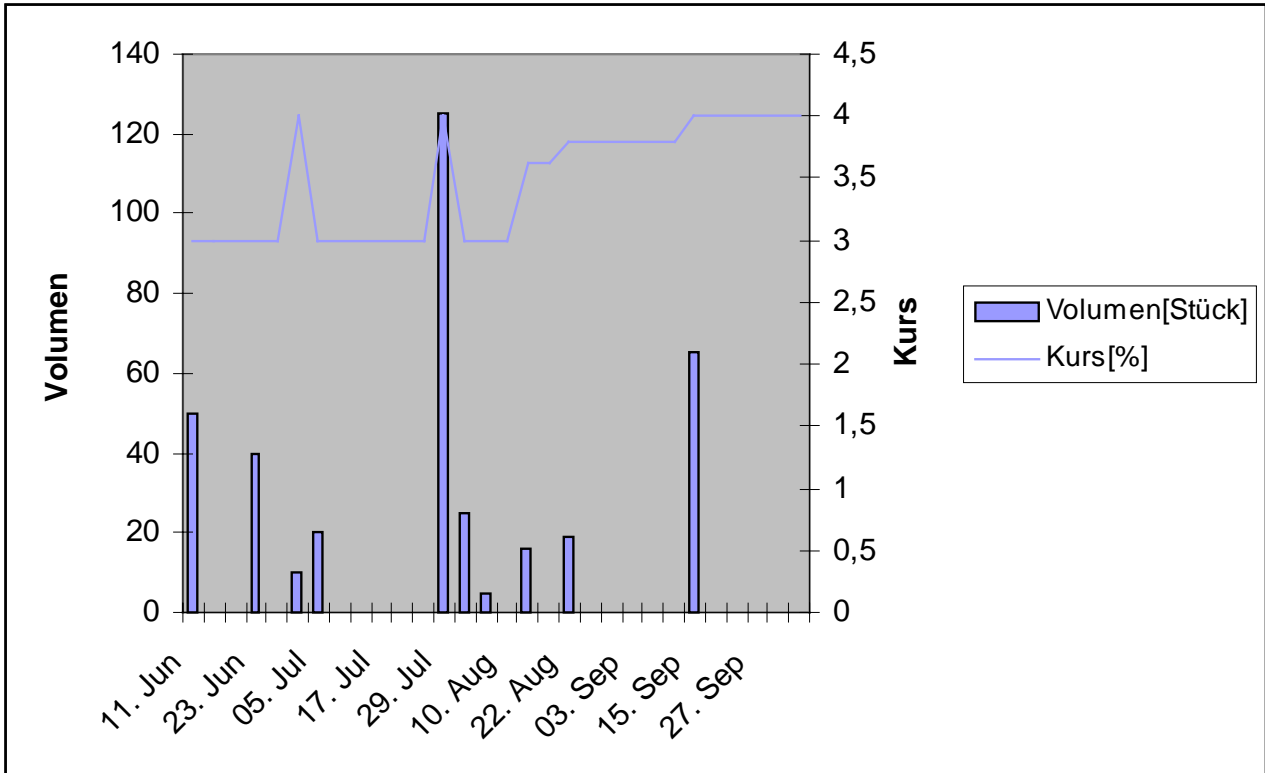
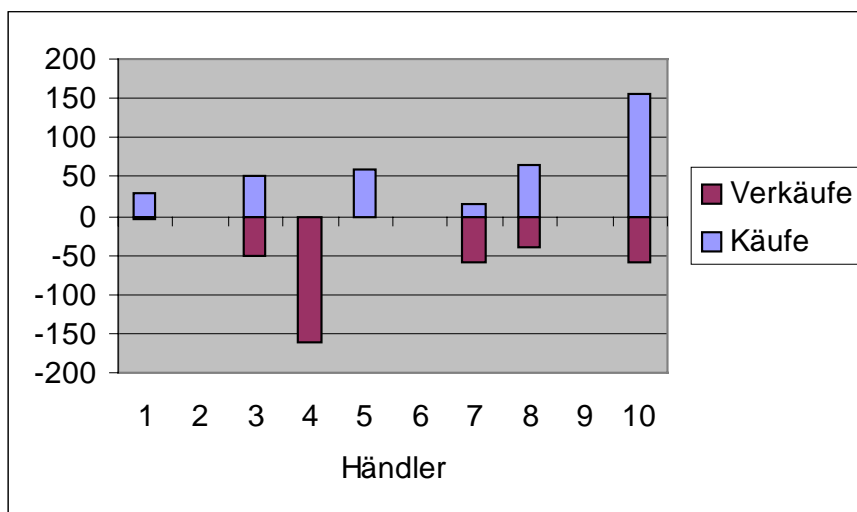


Abbildung 2.32: Gehandelte Aktien (Käufe und Verkäufe) – LiF



Der Kurs der LiF-Aktie schwankte zwischen 3 und 4%, und schloß mit 4%. Sie war damit um fast 2% überbewertet.

Gehandelt wurde sie, im Vergleich zu den anderen Aktien, relativ gleichmäßig, auch die Transaktionen Ende Juli, Anfang August waren nicht so stark ausgeprägt. Die letzten Transaktionen fanden Mitte September statt, sie bestimmten auch den Schlußkurs.

Die Dominanz der Trader 4 und 10 ist zwar vorhanden, aber auch nicht so stark wie in anderen Märkten.

Negativ wirkte sich auf die LiF-Aktie sicher aus, daß sie, wie schon erwähnt, nicht mit einer Kommastelle gehandelt werden konnte.

Weiters ist zu bemerken, daß es das erste Antreten des LiF in Oberösterreich war, und deshalb das Ergebnis äußerst schwer vorherzusagen war.

## **II.5 Zusammenfassung**

Der Oberösterreichische Landtagswahlmarkt ist ein schlechtes Beispiel für einen Political Stock Market. Die geringe Teilnehmerzahl, deren Zusammensetzung, die Dominanz einzelner Trader in den Märkten und vor allem das geringe Interesse haben das Ergebnis natürlich sehr negativ beeinflusst.

Ein wesentlicher Grund für das teilweise schlechtere Abschneiden von APSM gegenüber den amerikanischen Projekten kann in der höheren Anzahl der vorauszusagenden Variablen liegen. In Untersuchungen in den Jahren 1990-1994, die amerikanische PSMs untereinander verglichen, konnte festgestellt werden, daß die Prognosegenauigkeit mit Anzahl von handelbaren Aktien abnimmt.<sup>24</sup> Das Ergebnis des APSM 97 liegt aber noch wesentlich unter dem vorangegangener österreichischer Experimente.

Die Ergebnisse der bisherigen Landtagswahlprojekte (Steiermark, OÖ) schneiden im Vergleich zum Ergebnis des APSM zur Wien-Wahl 96 sehr schlecht ab. In Zukunft muß es gelingen, mehr Teilnehmer aus dem entsprechenden Bundesland zu motivieren, am Projekt teilzunehmen, ansonsten werden die Ergebnisse auch in Zukunft unbefriedigend sein.

Eine große Teilnehmerzahl ist erforderlich, um stark ausgeprägte Teilnehmerkreise (z.B. nur männliche Teilnehmer), und die Dominanz von wenigen Teilnehmern zu verhindern.

Dadurch wird auch der negative Einfluß von Manipulationsversuchen von wenigen Teilnehmern reduziert.

Es muß also versucht werden, Anreizsystem zu schaffen, um mehr Menschen zur Teilnahme zu bewegen und diese zu motivieren, gut am Markt abzuschneiden. Denn

---

<sup>24</sup> Vgl. [Berg, Forsythe, Rietz 1996]

nur so werden Teilnehmer angehalten, sich private Informationen zu beschaffen, um einen Wettbewerbsvorteil gegenüber anderen Teilnehmern zu haben.

Ein größerer Teilnehmerkreis kann bei größerem medialen Interesse automatisch erreicht werden. Außerdem wird sich die Zahl der Internetuser in Österreich in den nächsten Jahre weiter verstärken, der potentielle Teilnehmerkreis wird also größer.

Zusätzlich muß aber verstärkt für derartige Projekte geworben werden, sei es durch Aushänge an Universitäten, Behörden,.. oder durch Einschaltungen in Zeitungen, Zeitschriften,..

Um die Teilnehmer mehr zu motivieren könnte das Projekt in jener Form ablaufen, wie auch Aktienspiele durchgeführt werden.

Das heißt, daß für gut handelnde Teilnehmer ein Bonus in Aussicht gestellt wird. Wenn es z.B. Preise für die besten Teilnehmer gibt, steigt die Motivation, gut abzuschneiden. Als Beurteilungskriterium für eine erfolgreiche Teilnahme kann die Rendite<sup>25</sup> eines Teilnehmers herangezogen werden.

Weiters darf auf die neue Software gespannt sein, da die zur Zeit verwendete einige Nachteile aufweist. Auch hier liegen noch Verbesserungspotentiale.

---

<sup>25</sup> Rendite =  $(\text{Endkapital}/\text{Ausgangskapital}-1)*100\%$

### **III. Prognosen durch Meinungsumfragen**

#### **III.1 Umfrageforschung in der empirischen Sozialforschung**

##### **III.1.1 Methoden der empirischen Sozialforschung**

Innerhalb der Methoden der empirischen Sozialforschung nimmt die Umfrageforschung einen wichtigen Platz ein – sie dürfte im öffentlichen Bewußtsein überhaupt mit Sozialforschung gleichgesetzt werden.

In der empirischen Sozialforschung will man das overte (soziale) Verhalten und Handeln von Personen sowie ihre Haltungen, Meinungen und Einstellungen erfassen. Will man in Erfahrung bringen, wie sich Personen verhalten oder verhalten haben oder wie ihre Einstellungen zu einer bestimmten Frage sind, so bieten sich fünf Möglichkeiten an.<sup>26</sup>

- Befragung von Personen – Umfrageforschung
- Gruppendiskussionen
- Beobachtung des Verhaltens von Personen
- Durchführung von Experimenten
- Inhaltsanalyse

Die Umfrageforschung, die Gegenstand der folgenden Betrachtungen ist, kann man weiter einteilen, so daß sich ein breites Spektrum an Möglichkeiten und Methoden<sup>27</sup> ergibt, innerhalb der Forschungspraxis haben sich aber die Methoden laut Abbildung 3.1 eingespielt.

---

<sup>26</sup> vgl. [Friedrichs 1980] S. 192

<sup>27</sup> „Eine Methode ist ein spezielles System von Regeln, das die Tätigkeit bei der Erlangung neuer Erkenntnisse und der praktischen Umgestaltung der Wirklichkeit organisiert“ [Böhnisch 1972] S. 21 zitiert in [Friedrichs 1980] S. 189



Abbildung 3.1: In der Forschungspraxis bewährte Methoden der Umfrageforschung<sup>28</sup>

	Fragenwortlaut gleich/nicht gleich	Erhebungseinheit Einzelperson/Personengruppe	Medium schriftlich/mündlich
eine Person	"Einzelfallstudie"		
mehrere Personen	Tiefeninterviews (nicht standardisiertes Einzelinterview)  "Gruppengespräch", "Gruppenbefragung"		
viele Personen ("Stichprobe")	Repräsentativerhebung (standardisiertes Einzelinterview, mündlich)  "Schriftliche Befragung"		
alle Personen	"Totalerhebung"		

Bei Meinungsumfragen zu Wahlprognosen werden „Repräsentativumfragen“, „Polls“, also standardisierte, strukturierte Interviews mit Einzelpersonen, schriftlich, mündlich, telefonisch, über Internet,... durchgeführt.

### **III.1.2 Prognosen und Hypothesenbildung**

Prognosen versuchen, eine erklärende Sicht der Dinge abzugeben. Hat man in der Naturwissenschaft oder Physik ein Prognose abzugeben, so kann man dies bei Kenntnis der Gesetzmäßigkeiten relativ einfach durchführen. Hält man einen Stein in der Hand, und kennt die Randbedingung „Wir stehen auf der Erde“, so kann man unter der Heranziehung der Gesetzeslage „Die Gravitationskraft wirkt auf alle Körper“ die Prognose „Der Stein wird beim Auslassen auf die Erde fallen“ mit 100%iger Sicherheit voraussagen.

Die Schwierigkeit von Prognosen in der Sozialwissenschaft besteht darin, daß es Gesetze dieses Typs nicht gibt, daß zur Prognose also gesetzesähnliche, jedenfalls

<sup>28</sup> vgl. [Karmasin/Karmasin 1977] S. 12

bestmöglich bestätigte Hypothesen herangezogen werden müssen, so daß sich ein mehr oder wenig großer Vagheitsspielraum ergibt.

Zur Hypothesenbildung und Annahme oder Verwerfung dieser, werden die Methoden der Statistik angewandt, auf dies wird später noch näher eingegangen.

### **III.1.3 Trenduntersuchungen**

Eine Besonderheit, die Meinungsumfragen zu Wahlen darstellen, ist, daß sie sogenannte Trenduntersuchungen sind, das heißt, daß Personengruppen zu einem identischen Fragenprogramm mehrmals befragt werden.

Während bei Paneluntersuchungen die Personengruppe gleich bleibt, wird darauf bei Trenduntersuchungen nicht Wert gelegt. Man kann also nur feststellen, ob insgesamt eine Veränderung stattgefunden hat und nicht, ob und welche Einzelpersonen ihre Meinung oder ihr Verhalten geändert haben.

Auch die Feststellung von Änderungsquoten bei Trenduntersuchungen kann problematisch sein. So wird z.B. bei einem Wechsel von 20% vom Lager A ins Lager B, und 20% vom Lager B ins Lager A, leicht „alles ist gleichgeblieben“ interpretiert, obwohl sehr wohl eine Änderung stattgefunden hat. Dieses Manko kann man durch gezielte Befragung einigermaßen ausgleichen, dazu mehr in Kapitel III.3.

Wichtig bei Trenduntersuchungen ist auch die Feststellung, ob gewisse Verschiebungen signifikant sind oder nicht. Nachdem Untersuchungsergebnisse statistischen Schwankungsbreiten unterliegen, kann bei einer Änderung innerhalb dieser Schwankungsbreite nicht von einer signifikanten Änderung gesprochen werden. Auch hier bedient man sich statistisch etablierter Tests.

Ziel der Trenduntersuchungen ist es, eine gewisse Dynamik in der Meinungsbildung herauszufinden, auch dann, wenn die Änderungen innerhalb der Schwankungsbreite liegen, dies erfordert viel Erfahrung und Gespür.

*„Ein Wahlkampf entwickelt in der letzten Woche eine spezielle Dynamik. Hier bewegt sich noch sehr viel. Beobachtete Kurven nach unten oder oben verstärken sich meist noch.“<sup>29</sup>*

## **III.2 Die Stichprobenauswahl**

### **III.2.1 Grundgesamtheit**

Unter Grundgesamtheit versteht man die Menge aller gleichartigen Individuen oder Objekte, auf die sich die Beobachtung beziehen soll. Die Summe aller Merkmalsträger bilden die Grundgesamtheit, häufig wird auch der Begriff „Population“ verwendet.

Die Grundgesamtheit muß natürlich je nach Aufgabenstellung definiert werden, eine eindeutige Definition ist für die Qualität einer Studie von großer Bedeutung. In Falle von politischen Wahlen bilden alle wahlberechtigten Staatsbürger Österreichs die Grundgesamtheit, es ist aber notwendig, die genaue Struktur bezüglich Alter, Bildung,... zu kennen, um eine gezielte Planung der Stichprobe vornehmen zu können.

Die Struktur der Grundgesamtheit kann man in Österreich aus Wählerverzeichnissen und Volkszählungen entnehmen.

Um festzustellen, wie häufig ein bestimmtes Merkmal in einer Population vorkommt, kann man natürlich auch die gesamte Population in die Untersuchung einbeziehen, dies ist natürlich nur bei relativ kleinen Grundgesamtheiten möglich, wie etwa bei einer Untersuchung auf betrieblicher Ebene.

---

<sup>29</sup> Georg Michenthaler von IFES im Interview vom 28.10.97

Die Stichprobenerhebung hat aber den enormen Vorteil, daß dem einzelnen Fall unvergleichlich mehr Aufmerksamkeit geschenkt werden kann, als dies bei einer Vollerhebung möglich ist. Dadurch kann sich der sachliche Fehler der Untersuchung soweit verringern, daß dies den Fehler durch die Schwankungsbreite einer statistischen Auswahl ausgleicht oder übersteigt.<sup>30</sup>

*„Unter einer Stichprobenerhebung wird eine nach statistischen Kriterien angelegte Untersuchung verstanden, bei der eine repräsentative Auswahl stellvertretend für die Grundgesamtheit beobachtet wird.“<sup>31</sup>*

Ziel der Stichprobe ist, einerseits Hypothesen zu entwickeln und anhand der Ergebnisse Generalisierungen (Verallgemeinerungen) von der Stichprobe auf die Grundgesamtheit vorzunehmen (Repräsentationsschluß), andererseits Hypothesen anhand der Stichprobe zu testen.

Voraussetzung der Stichprobe sind:

- Die Stichprobe muß ein verkleinertes Bild der Grundgesamtheit hinsichtlich der Heterogenität der Elemente und hinsichtlich der Repräsentativität der für die Hypothesenbildung relevanten Variablen sein.
- Die Einheiten oder Elemente der Stichprobe müssen definiert sein.
- Die Grundgesamtheit sollte angebbar und empirisch definiert sein. (Ist im Falle von Meinungsumfragen zu Wahlen in Österreich gegeben)
- Das Auswahlverfahren muß angebbar sein und die erste Forderung erfüllen, und die Auswahl der Stichprobe **zufällig** sein. Damit sind nicht die Daten in der Stichprobe gemeint, sondern das Verfahren bei der Ziehung der Stichprobe. Auf dieses wichtige Merkmal wird bei der Vorstellung der Auswahlverfahren noch näher eingegangen.

---

<sup>30</sup> vgl. [Scheuch 1976] S. 310 zitiert in [Karmasin/Karmasin 1977] S. 223

<sup>31</sup> vgl. [Karmasin/Karmasin 1977] S. 225

### **III.2.2 Statistische Grundkenntnisse**

Statistische Verteilungen werden im einfachen Fall mit einem Lageparameter und einem Streuungsparameter definiert.

Arithmetische Mittelwert:  $\mu = 1/n * \sum x_i$

Varianz:  $\sigma^2 = 1/(n-1) * \sum (x_i - \mu)^2$

-> Standardabweichung  $\sigma$

Die am häufigsten anzutreffende empirische Verteilung ist die symmetrische Verteilung. Die Binomial- und hypergeometrische Verteilung haben vor allem für kleine Stichproben Bedeutung. Mit zunehmender Anzahl der Fälle nähern sie sich der Normalverteilung an, die natürlich theoretisch erst ab einer unendlich großen Stichprobe gewährleistet ist. In der Praxis wird aber in der Umfrageforschung immer die Normalverteilung als Maßstab herangezogen, da sie bei den ausgewählten Stichproben hinreichend genau approximiert wird.

Abbildung 3.2: Normalverteilung

Die Wahrscheinlichkeit, daß ein gewisser Wert in einem Bereich liegt, wird *Signifikanzniveau* genannt.

Laut Abbildung 3.2 ist das Signifikanzniveau für  $\mu - \sigma < x < \mu + \sigma$  68,3%, für  $\mu - 2\sigma < x < \mu + 2\sigma$  95,5% und für  $\mu - 3\sigma < x < \mu + 3\sigma$  99,7%.

In der Umfrageforschung wird mit der doppelten Standardabweichung gerechnet, die 95,5% aller möglichen Stichproben umfaßt.

Da bei Umfragen oft keine quantitativen Werte gemessen werden, sondern, wie z.B. auch bei der „Sonntagsfrage“, „Merkmalsanteile“, muß man die statistischen Parameter anders definieren. Angenommen ein Merkmal kommt mit einer Wahrscheinlichkeit  $p$  in einer Aussage einer Stichprobe  $n$  vor, so ergibt sich der arithmetische Mittelwert logischerweise mit  $\mu = n \cdot p$ .

Die Varianz, die durchschnittliche quadratische Abweichung vom Mittelwert, ist dann  $\sigma^2 = n \cdot p \cdot (1-p)$ . Bringt man noch die Endlichkeit von Grundgesamtheiten durch den sogenannten Endlichkeitsfaktor ins Spiel und will die Standardabweichung als Prozentergebnis, so ergibt sich die Formel zu:<sup>32</sup>

$$\sigma = \sqrt{\frac{p \cdot (100-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}} \quad [\%]$$

Der Endlichkeitsfaktor geht bei großen Grundgesamtheiten sehr schnell gegen 1, z.B. bei der Oberösterreichischen Landtagswahl mit  $N = 973.750$  Wahlberechtigten<sup>33</sup> und bei Stichprobengrößen von  $600$ <sup>34</sup>, ergibt sich ein Endlichkeitsfaktor von  $1,0000005$ .

Zur Verdeutlichung des Einflusses des Merkmalsanteil folgendes Beispiel:

Bei einer repräsentativen Stichprobe werden 1000 Autofahrer befragt, ob sie Gurte anlegen oder nicht. Das Ergebnis ist z.B. 80% legen den Gurt an, 20% nicht oder fast nie. Die Standardabweichung ist dann unter Vernachlässigung des

---

<sup>32</sup> vgl. [Karmasin/Karmasin 1977] S. 232

<sup>33</sup> vgl. Endergebnis der OÖLTW97 des Statistischen Dienstes des Amtes der öö Landesregierung

Endlichkeitsfaktors 1,26%. Der schlechteste Fall, also die höchste Standardabweichung (1,58%) liegt bei einem Merkmalsanteil von 50%:50%.

Aufgrund dieser Überlegungen ist es relativ leicht, die notwendige Größe einer Stichprobe zu berechnen. Man muß die Formel für die Standardabweichung nur auf  $n$  umformen.

Die zulässige Schwankungsbreite ist vom Forscher anzugeben, die Merkmalsausprägung muß geschätzt werden, kann aber mit 50% für den schlechtesten Fall angenommen werden.

Beispiel: Bei den „SPECTRA“-Befragungen zur OÖLTW wurden 500, 1000 oder 1500 Personen befragt. Die Stichprobenschwankungen für einen 95,5%-Bereich errechnen sich laut obiger Formel ( $p = 50\%$ , Endlichkeitsfaktor = 1):

$$n = 500 \Rightarrow 2\sigma = \pm 4,5\%$$

$$n = 1000 \Rightarrow 2\sigma = \pm 3,2\%$$

$$n = 1500 \Rightarrow 2\sigma = \pm 2,6\%$$

Dies sind auch die von „SPECTRA“ angegebenen Schwankungsbreiten.<sup>35</sup>

Diese Berechnungsformel gilt aber nur für den einfachsten Fall, das total zufällige Auswählen aus einer Grundgesamtheit. Dies ist z.B. der Fall bei  $N$  Kugeln in einem Gefäß, wobei eine gewisse Zahl davon schwarz ist. Man kann dann durch Ziehen einer Stichprobe  $n$  laut obiger Formel vorgehen.

Die Randbedingungen in der Umfrageforschung sind natürlich andere, und bei den in der Praxis üblichen Auswahlverfahren kommen noch andere Einflüsse hinzu, die das Ergebnis positiv oder negativ beeinflussen können. Auf dies wird später bei den einzelnen Auswahlverfahren noch näher eingegangen.

---

<sup>34</sup> vgl. Studie von IFES eine Woche vor der Landtagswahl

<sup>35</sup> vgl. [Watzl 1997]S. 10

### III.2.3 Strukturierung des Objektbereiches

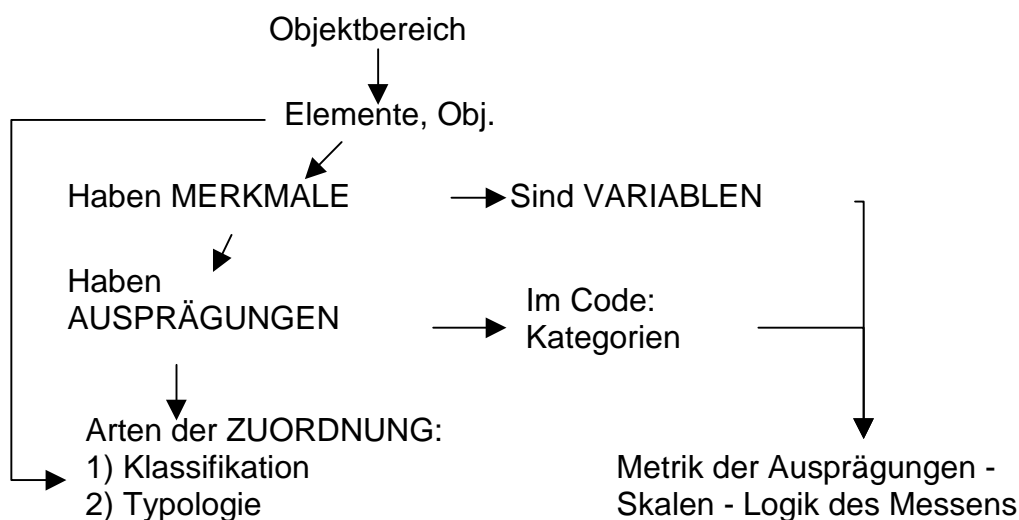
#### a) Objektbereich, Merkmal & Variable

Die theoretisch am einfachsten wirkende Auswahlmethode, die einfache Zufallsauswahl bringt in der Praxis oft Schwierigkeiten mit sich. Sie setzt voraus, daß von der Grundgesamtheit eine vollständige Adressenkartei vorliegt, in der jedes Mitglied genau einmal vertreten ist, aus der dann rein zufällig soviele Adressen gezogen werden, wie in der Stichprobe enthalten sein sollen. Dazu ist es nötig, daß die Kartei vollständig und für den Zeitpunkt der Untersuchung aktuell ist.

Einige übliche Auswahlverfahren entsprechen keineswegs diesen Voraussetzungen. Es ist daher notwendig, die zugrundeliegende Grundgesamtheit zu analysieren und strukturieren, um geeignete Stichproben ziehen zu können. Selbst bei Auswahlverfahren mit Adressenlisten ist dies notwendig, um zufällig auftretende Ausfälle durch Gewichtung ausgleichen zu können.

Ausgangspunkt bei der Strukturierung des Objektbereichs ist die Unterscheidung von Objektbereich, Elemente und Merkmalen.

Abbildung 3.3: Elemente der Strukturierung eines Objektbereichs





In einem Objektbereich (z.B. Stadt) gibt es Elemente bzw. Objekte (Einwohner), diesen werden Eigenschaften (Merkmale) zugeschrieben. Elemente können nie in ihrer Totalität gemessen werden, sondern immer nur einige Merkmale dieser Elemente. Ein Merkmal, daß meßbar ist und verschiedene Werte, und zwar mindestens zwei, annehmen kann, wird als Variable bezeichnet.

Eine Typologie von Variablen läßt sich nach der Art der Merkmalsausprägung, die sie erfassen, erstellen. Demnach unterscheidet man qualitative und quantitative Variablen.

Bei quantitative Variablen erfolgt die Zuschreibung der Merkmalswerte relativ einfach über Maßeinheiten. Ein Beispiel bei Umfragen zu Wahlergebnissen ist die quantitative Variable *Alter*.

Die überwiegende Anzahl von Variablen in der empirischen Sozialforschung sind aber qualitative. Sie können in dichotomer Ausprägung auftreten (nur zwei Merkmalswerte möglich), wie z.B. *Geschlecht*, oder in mehreren Klassen, wie *Schulbildung*, *Tätigkeit*, *Wohnregion*.

Wie in Abbildung 3.3 angedeutet, kann die Zuordnung der Elemente und Variablen durch Klassifikation oder Typologie erfolgen, wobei die vollständige Zuordnung der Klassifikation bei Meinungsumfragen zu Wahlprognosen vorherrscht.

#### b) Klassifikation

Bei der Strukturierung des Objektbereichs wird versucht, eine gewisse Ordnung in die beobachteten Erscheinungen zu bringen. Dabei kann es sich um Personen, Normen,... handeln. Einer Menge von Elementen wird eine Menge von Merkmalen zugeordnet, man nimmt eine Klassifikation vor.

Der Umfang der Klassifikation hängt von der Menge der Merkmale und ihrer Dimensionen ab. Zahlreiche Objekte werden gleichzeitig definiert, die zumeist

aufgrund empirisch gefundener Zusammenhänge entsteht, also statistischer Hypothesen. Erfahrungen zeigten, daß die Wahlentscheidung von Personen im wesentlichen von den Merkmalen Geschlecht, Alter, Bildung, Tätigkeit abhängt.

Eine Klassifikation muß drei Anforderungen genügen:<sup>36</sup>

- Eindeutigkeit: Jedem Element kann die Merkmalsausprägung eines Merkmals zugeschrieben werden.
- Ausschließlichkeit: Nur eine Ausprägung des Merkmals trifft zu.
- Vollständigkeit: Kein Element ist ohne einer Merkmalsausprägung

Die Erstellung von sinnvollen Maßeinheiten, die eine Erfassung von einzelnen Elementen gestattet, und die eine größtmögliche statistische Manipulation der Daten erlauben (nachträgliche Gewichtung) setzen hohe theoretische und empirische Kenntnisse voraus.

Wichtig für die Messung der Variablen ist die *Eindimensionalität*. Während im Bereich der Naturwissenschaften meist eine Variable einfach isoliert gemessen werden kann, ist dies im Bereich der Sozialwissenschaft nicht ohne weiteres möglich. Der Aufbau des Interviews und die Fragestellung haben daher wesentlichen Einfluß auf die Qualität der Meßergebnisse (siehe auch Kapitel III.3).

Die am häufigsten verwendete *Skalenform* bei der Messung von Variablen bei Meinungsumfragen zu Wahlergebnissen ist die *Nominalskala*. Erfasst werden qualitative Variablen, die Merkmalsausprägungen werden über klassifikatorische Begriffe zugeordnet. Diese werden dann mit Zahlen versehen, z.B. 1 = Selbständige, 2 = Arbeiter, 3 = Angestellte,...

Diese Zahlen haben den Status von Kennziffern und können beliebig vertauscht werden. Bei der Auswertung werden dann die absoluten oder relativen Häufigkeiten der Merkmale beachtet.

---

<sup>36</sup> vgl. [Friedrichs 1980] S. 88,89

### c) Typologie

Bei der Erstellung der Typologien interessiert weniger der genaue Ausprägungsgrad der Variablen, sondern vielmehr die Selektion und die Zusammenhänge von Variablen, die zu einer sinnvollen Gliederung des Objektbereichs dienen können.

Die Einteilung von Pflanzen ist z.B. eine Typologie oder die Einteilung von Personen in der Psychologie.<sup>37</sup>

Es wird versucht, eine Ordnung einzuführen, die es erlaubt, die einzelnen Elemente eindeutig voneinander abzuheben und sie gleichzeitig zu sinnvollen Merkmalsklassen zusammenzufassen. Die Zuordnung der Elemente zu Klassen kann unvollständig sein.

Tabelle 3.1: Klassisches Beispiel einer Typologie, die Formen des Verhaltens von *Merton*:<sup>38</sup> (Ein „Plus“ bedeutet Zustimmung, ein „Minus“ Ablehnung, ein „Plus und Minus“ Ablehnung der vorherrschenden Ziele und Substitution neuer Ziele)

Formen der Anpassung	Kulturelle Ziele	Institutionalisierte Mittel
I. Konformität	+	+
II. Innovation	+	-
III. Ritualismus	-	+
IV. Rückzug	-	-
V. Rebellion	+/-	+/-

### **III.2.4 Auswahlverfahren**

Im Rahmen dieses Abschnitts wird hauptsächlich auf die in der Praxis tatsächlich zur Anwendung gelangenden Auswahlverfahren eingegangen. Die Kennzeichen aller Verfahren sind:<sup>39</sup>

---

<sup>37</sup> Vgl. [Karmasin/Karmasin 1977] S. 131

<sup>38</sup> Vgl. [Merten 1957] S. 139ff, in [Friedrichs 1980] S. 90

<sup>39</sup> Vgl. [Friedrichs 1980] S. 130

- Die Auswahl wird entweder so vorgenommen, daß jedes Element die gleichen Chancen hat gezogen zu werden (Wahrscheinlichkeitsauswahl), oder daß Elemente mit bestimmten Merkmalen bevorzugt werden (bewußte Auswahl).
- Stichprobeneinheiten können Elemente oder Cluster von Elementen sein. Die Auswahl kann sich daher von Elementen aus der Grundgesamtheit beziehen (einstufig). Bei der mehrstufigen Auswahl werden zuerst Cluster, und in diesen Cluster die Elemente oder wiederum Cluster ausgewählt.

Abbildung 3.4: Grundgesamtheit und Stichprobeneinheit <sup>40</sup>

Elemente	Merkmale									
	M1	M2	M3	M4	M5	M6	M7	...	Mj	
	a	b	c							
E1	x			A						
E2	x									I
E3	x	B		C						
E4	x									
E5	x									
E6	x									II
E7	x									
...										
Ei		x	D							III
		x								
		x								

a,b,c = Ausprägungen des Merkmals M1

I,II,III = Schichten

A,B,C,D = Cluster

---

<sup>40</sup> Vgl. [Friedrichs 1980] S. 131

a) Die reine Zufallsauswahl und das systematische Sample

Bei der reinen Zufallsauswahl enthält jedes Mitglied in der Grundgesamtheit, also jede Adresse in der Kartei, eine laufende Nummer und die Adressenauswahl geschieht dann durch Zufallszahlen.

Im Gegensatz zu einer reinen Zufallsauswahl werden beim systematischen Sample die Adressen aus einer Kartei nicht durch Zufallszahlen ermittelt, „*sondern man zieht systematisch jede soundsovielte*“.<sup>41</sup>

Der Abstand zwischen den gezogenen Adressen berechnet sich mit  $k = N/n$ . Bei  $N = 973.750$  Wahlberechtigten und einer Stichprobe  $n = 705$  würde jede  $k = 1381$ -te Adresse gezogen werden. Die Zufälligkeit der Auswahl wird dadurch gewährleistet, daß ein „zufälliger“ Anfang gemacht wird. Das heißt, daß die erste Adresse mit Hilfe einer Zufallszahl aus den ersten  $k$  Adressen, in diesem Beispiel zwischen 1 und 1381, gezogen wird. Ab dann wird jede  $k$ -te Adresse gezogen.

Systematisches Sampling ist einfach und praktisch zu handhaben, allerdings muß der rein zufällige Aufbau der Kartei gegeben sein. Normalerweise sind Karteien nach gewissen Kriterien sortiert, wie alphabetische Reihenfolge oder Kundennummern. Wenn eine Kartei rein zufällig angeordnet ist, entspricht die systematische Auswahl der zufälligen und die Genauigkeit ist die gleiche.

Dieses Verfahren wird allerdings nur bei relativ kleinen und örtlich nahegelegenen Grundgesamtheiten durchgeführt. Bei Befragungen zu Wahlprognosen ist dies meist nicht der Fall, die Kosten für Fahrtspesen wären zu hoch, es wird daher auf andere Auswahlverfahren zurückgegriffen.

---

<sup>41</sup> Vgl. [Holm 1982] S. 146

## b) Stratifizierte Stichproben

Für das Abgehen vom systematischen Sample gibt es eine Reihe von Gründen, die sich aus theoretischen Überlegungen herleiten. Die Elemente der Grundgesamtheit können nach verschiedenen Kriterien in Unterklassen zusammengefaßt werden (z.B. Region) und in diesen Unterklassen verteilen sich verschiedene Merkmale unterschiedlich auf die einzelnen Klassen.

Man spricht von Stratifizieren oder Schichten, wenn man die Grundgesamtheit auf bestimmte Gruppierungsmerkmale hin untersucht und diese bei der Anlage der Stichprobe berücksichtigt. Die einzelnen Bevölkerungsschichten wie Geschlecht, Altersgruppen, regionale Gliederungen,... werden in ihrem richtige Verhältnis im Sample repräsentiert, die Zufälligkeit der Auswahl in diesem Sample muß natürlich wieder gewährleistet sein.

Je mehr Schichten man bildet, umso mehr wird sich der Fehler des Auswahlverfahrens verringern. Die Kosten dürfen aber nicht zu sehr in die Höhe schnellen, die Schichtung ist also ein Optimum aus Auswahlgüte und Auswahlkosten.

In vielen Fällen weist eine geschichtete Auswahl einen kleineren Zufallsfehler auf, dieser Präzisionsgewinn läßt sich leicht veranschaulichen, wenn man die Varianzzerlegung für ein beliebiges Merkmal  $y$  betrachtet.<sup>42</sup>

$$\sigma(y) = 1/N * \sum [N_i * \sigma(y_i)] + \sigma(y_z)$$

$\sigma(y)$  .....Gesamtvarianz in der Auswahlpopulation

$N$  .....Anzahl der Elemente in der Auswahlpopulation

$N_i$  .....Anzahl der Elemente in der  $i$ -ten Schicht ( $i=1.....K$  Schichten)

$\sigma(y_i)$  .....Varianz innerhalb der  $i$ -ten Schicht (Binnenvarianz)

$\sigma(y_z)$  ....Varianz zwischen den Schichten

Die gesamte Merkmalsvarianz setzt sich also zusammen aus der Varianz zwischen den Schichten und der Varianz innerhalb der Schichten. Die Reduktion des

statistischen Zufallsfehlers, der Schichtungseffekt ergibt sich mit dem Stichprobenumfang  $n$ :

$$\text{Schichtungseffekt} = 1/n * \sigma(y_z)$$

Der Schichtungseffekt ist im schlechtesten Fall null, nämlich dann, wenn die Varianz zwischen den Schichten gleich null ist. Der Schichtungsgewinn ist umso größer, je stärker die Abweichung der Schichtenmittelwerte voneinander ist, bzw. je größer die Varianz zwischen den Schichten ist.

Es wird also versucht, möglichst solche Schichtungen auszuwählen, die in sich sehr homogen, untereinander aber sehr verschieden sind. Man wird also nach Merkmalen schichten, die mit dem Untersuchungsziel eng zusammenhängen, im Falle von Meinungsumfragen zu Wahlprognosen ist dies z.B. das Merkmal regionale Herkunft.

Bei der Schichtung unterscheidet man die proportionale und die disproportionale Stichprobe. Bei der proportionalen entspricht der Anteil der Stichproben in der Schicht am Gesamtsample dem Anteil dieser Schicht an der Gesamtpopulation.

Disproportional geschichtete Auswahlen werden vor allem dann bevorzugt, wenn für die Analyse wichtige Teilstichproben unterrepräsentiert sind.

Bei einer repräsentativen Umfrage bei der OÖLTW würde eine Schichtung nach Regionen z.B. bedeuten, daß bei einem Stichprobenumfang von 705 Interviews 171 auf das Hausruckviertel entfallen, es ist also möglich über das Hausruckviertel statistisch abgesicherte Aussagen zu machen.

Auf das Innviertel würden aber nur 107 Interviews entfallen, es wären keine statistisch abgesicherten Aussagen möglich. Geht man von einer Mindestgröße von 150 Interviews pro Region aus, so ergebe dies bei einer proportionalen Erhöhung 240 Interviews für das Hausruckviertel und 989 für ganz Oberösterreich.

---

<sup>42</sup> Vgl. [Költringer 1993] S. 102,103

Um diese erheblichen Mehrkosten hintanzuhalten, wird mit disproportionalen Samplen gearbeitet. Die überrepräsentierten Vierteln müssen dann natürlich entsprechend durch Gewichtung nachbearbeitet werden.

### c) Mehrstufiges Auswahlverfahren und Clusterstichprobe

Bei großen Stichproben begnügen sich manche Institute nicht mit einer Vorschichtung. Sie teilen die Grundgesamtheit in Untereinheiten (Samplepoints) ein. Bei einem dreistufigen Verfahren wird in der ersten Ziehungsstufe eine entsprechende Zahl von Samplepoints per Zufall ausgewählt, in der zweiten Ziehungsstufe wird eine entsprechende Haushaltsadresse, in diesem Haushalt die entsprechende Zielperson per Zufall gezogen.

Wird bei mehrstufigen Auswahlverfahren, die Auswahl auf Merkmale bezogen, die in sich verhältnismäßig homogen sind, so spricht man von einer Clusterstichprobe. In der Praxis sind mehrstufige Auswahlverfahren immer Clusterstichproben, da man die Fahrtspesen der Interviewer deutlich reduzieren kann. Außerdem stehen meist nur Haushaltsadressen zur Verfügung hat, in einem Haushalt wird aber nur eine Person befragt.

Beim Auswählen von Cluster läuft man Gefahr, daß ganze Klassen von Einheiten mit bestimmten Merkmalen ein- oder ausgeschlossen werden, dies drückt sich in einem höheren Stichprobenfehler aus, der aber akzeptiert wird. *„Der Verlust der Stichprobengüte steht in keinem Verhältnis zu den eingesparten Kosten“*<sup>43</sup>

Der negative Effekt der Clusterauswahl auf die Stichprobengüte ist leicht verständlich. Wenn ein Merkmal in allen Klumpen gleich verteilt ist wie in der Auswahlpopulation, dann entspricht die Genauigkeit der Klumpenstichprobe jener der nicht geklumpten.

---

<sup>43</sup> Vgl. [Holm 1982] S. 153



Mit einer Verschlechterung muß aber gerechnet werden, wenn die Clusterelemente im Durchschnitt homogener sind als die Elemente der Auswahlpopulation. Je größer die Homogenität der Cluster, desto stärker hängt das Stichprobenergebnis von den jeweils ausgewählten Clustern ab und desto größer ist daher der statistische Zufallsfehler.

Der Effekt der Clusterauswahl auf die Stichprobenvarianz des Mittelwertes  $y^*$  der Variable  $y$  hängt von der Homogenität des Clusters (Korrelation  $\rho_h$ ) und der Anzahl der Stichprobenelemente je Cluster ab:<sup>44</sup>

$$\sigma(y^*) = [\sigma(y_i) / n] * [1 + (b-1) * \rho_h]$$

$\sigma(y^*)$  .....Stichprobenvarianz des Mittelwertes in Klumpenstichproben

$\sigma(y_i) / n$ .....Stichprobenvarianz des Mittelwertes in einfachen Zufallsstichproben

$\rho_h$  .....Homogenität des Cluster

$b$  .....Zahl der Stichprobenelemente je Cluster

$n$  .....Stichprobengröße

Beurteilt man den Präzisionsverlust mit der Maßzahl *Designeffekt* (= Stichprobenvarianz bei Verwendung von Clustering / Stichprobenvarianz bei einfacher Zufallsauswahl mit gleicher Stichprobengröße), so ergibt sich:

$$\text{Designeffekt} = 1 + (b-1) * \rho_h$$

Nur dann, wenn sich die Merkmalsverteilungen in den einzelnen Clustern nicht unterscheiden ( $\rho_h = 0$ ), entspricht die zu erwartende Präzision der Clusterstichprobe die der einfachen Zufallsstichprobe.

---

<sup>44</sup> Vgl. [Költringer 1993] S. 107

d) Die Flächenstichprobe (Random-Route-Verfahren)

Bei den bisher besprochenen Auswahlverfahren wurde davon ausgegangen, daß die Grundgesamtheit vollständig erfaßt ist und Adressen für die Auswahl zur Verfügung stehen. Dieses Material ist aber oft nicht zugänglich, vor allem in den USA verwendet man daher das Random-Route-Verfahren.

Bei diesem Verfahren wird das Untersuchungsgebiet gleich wie beim mehrstufigen Auswahlverfahren in gut abgegrenzte, annähernd gleich stark besetzte Flächeneinheiten (Cluster) aufgeteilt, die dann die Sampleeinheiten für die erste Auswahlstufe bilden. Voraussetzung ist allerdings, daß man die Größe der Flächen, also die Anzahl der Einwohner, kennt.

In den ausgewählten Clusters wird dem Interviewer dann ein Startpunkt vorgegeben, und er muß entlang einem vorgeschriebenem Weg die Haushalte und Zielpersonen nach einem vorgegebenem Zufallsverfahren auswählen, z.B. 2.Gasse, 3.Haus,...

Vom theoretischen Ansatz her entspricht dieses Verfahren dem eines mehrstufigen Samples. Es gibt aber zahlreiche praktische Einwände, vor allem die Zufallsauswahl wird angezweifelt. Die Interviewer sind schwer zu kontrollieren, und könnten unangenehme Viertel meiden und haben überdies noch zahlreich andere Möglichkeiten zur Manipulation. Untersuchungen haben gezeigt, daß der Grad der Manipulation sehr hoch werden kann, und daß vor allem erwerbstätige und männliche Personen bei der Auswahl nach Random-Route stark unterrepräsentiert sind.<sup>45</sup>

Aus diesen Gründen ist das Random-Route-Verfahren wirklich nur anzuwenden, wenn sich keinerlei Möglichkeit bietet, auf Adressenlisten der Grundgesamtheit zugreifen zu können.

---

<sup>45</sup> Vgl. [Költringer 1993] S. 115

### e) Das Quotenverfahren

Das Quotenverfahren ist eine zweite Möglichkeit, trotz fehlender Adressen zu einer brauchbaren Stichprobe zu kommen. Die Zielpersonen werden im Unterschied zum Random-Route-Verfahren nicht zufällig bestimmt, sondern der Interviewer bestimmt diese aufgrund von bestimmten Merkmalskombinationen selbst. Der Interviewer bekommt aber genaue Vorgaben, Quoten, wieviel Männer, Frauen, welche Altersgruppen, Berufsgruppen,... er befragen soll.

Es werden meist auch Schichtungen und Mehrstufen-Verfahren angewandt, die Quotenauswahl bezieht sich dann auf die letzte Stufe der Zielpersonenermittlung, auf das sampling innerhalb der Schichten.

Wichtig ist, daß man einem Interviewer nicht zu „schwere“ Quoten aufträgt. Das heißt, daß er Personen befragen muß, die in der Bevölkerung äußerst selten vorkommen, wie z.B. ein Bauer in der Großstadt, oder daß die Person einer Gruppe der Bevölkerung angehört, die aus Zeitnot sehr ungern ein Interview gibt, wie etwa Selbständige oder freiberuflich Tätige.

Da die Erfüllung der Quoten umso schwieriger wird, je mehr Quotenmerkmale herangezogen werden, beschränkt man sich meist auf die Merkmale Alter, Geschlecht, Bildung, und hofft, daß sich die übrigen Quotenmerkmale, die unter Umständen für die Befragung noch wichtig sein können, automatisch einstellen.

Die größte Kritik an der Quotenstichprobe ist, daß sich die Interviewer die Personen selbst aussuchen können. Der Interviewer trifft eine bewußte, oft subjektive Auswahl, allein schon deshalb, weil der Mensch kein Mechanismus ist, der reine Zufallsauswahl treffen kann. Die Interviewer sind schwer zu kontrollieren, greifen bei Probleme mit den Quoten auf ihren Bekanntenkreis zurück, diese werden irgendwann zu routinierten Antwortern. Es kommt auch häufig vor, daß sich der Interviewer von seinen Befragten zu anderen Personen empfehlen läßt, die in seine Quote passen, es kommt trotz der Quotierung zu clusterartigen Häufungen.

Obwohl die statistische Zufälligkeit der Auswahl keineswegs gegeben ist, hat sich das Quotenverfahren seinen Platz in der Praxis der Umfrageforschung, nicht zuletzt wegen seiner geringen Kosten und seinen durchführungstechnischen Vorteile gegenüber dem Random-Route-Verfahren, erkämpft.

Vergleichsuntersuchungen haben gezeigt, daß Quotenstichproben unter gewissen Voraussetzungen zu überraschend guten Ergebnissen führen. Dazu ist aber viel praktische Erfahrung im Umgang mit Quotenstichproben notwendig. Nicht nur die Quotierungsmerkmale selbst sind dabei wichtig, sondern auch die Kopplungen untereinander. Diese müssen so gewählt werden, daß der Interviewer nicht mehr die Chancen hat, nur „leichte“ Leute zu befragen. Die Anweisungen dürfen aber auch nicht zu schwer zu erfüllen sein, denn sonst wird der Interviewer überfordert und man ermutigt ihn, unsauber zu arbeiten.

### ***III.2.5 Nichterfassungsfehler, Ausfallsfehler und Gewichtung***

In Österreich werden das mehrstufige Verfahren und von kleineren Instituten auch das Quotenverfahren angewendet. Neben den bereits besprochenen, aufgrund statistischer Bedingungen auftretenden Fehlern, gibt es noch weitere Fehlermöglichkeiten, die in den Bereich der Stichprobenfehler fallen:

- Nichterfassungsfehler und
- Ausfallsfehler

Unter Nichterfassung versteht man den Umstand, daß Elemente der Grundgesamtheit keine Auswahlchance haben, weil sie in der Auswahlpopulation nicht vorhanden sind (z.B. Nicht-Telefonbesitzer in einer Telefonumfrage). Wenn die nichterfaßten Elemente andere Merkmale aufweisen als die erfaßten, dann wird das Ergebnis natürlich verzerrt.

Weitere Fehlerquellen können die Fehlerfassung (Existenz von Fremdelementen in der Auswahlpopulation, z.B. Firmenadressen bei Haushaltsbefragungen) und die

Übererfassung (Elemente der Grundgesamtheit kommen öfter in der Auswahlpopulation vor und haben daher eine erhöhte Auswahlwahrscheinlichkeit) sein.

Die größeren Meinungsforschungsinstitute in Österreich können auf ausgezeichnetes Adressenmaterial als Grundlage der Stichprobenziehung zurückgreifen, und zwar „öffentlich zugängliche Wählerlisten, welche die Gesamtheit der österreichischen Staatsbürger im Alter von über 18 Jahren nahezu fehlerfrei erfassen.“<sup>46</sup>

Übererfassungen sind dadurch ausgeschlossen, da nur Wählerlisten nationaler Wahlgänge herangezogen werden. Nichterfassungsfehler können aber auftreten, da die Adressenbestände meist alle vier Jahre, im Rhythmus der Nationalratswahlen, vollständig überarbeitet und aktualisiert werden. In der Zwischenzeit kann durch Gebäudeabtragungen, Umbauten, Umzüge und Todesfälle ein gewisser Anteil nicht mehr aktueller Adressen in der Kartei entstehen.

Ungültige Adressen, die sich aufgrund der Mobilität der Bevölkerung (Umzüge, Zuzüge) ergeben, sind im Prinzip kein Problem, da an einer Adresse ein Interview angestrebt wird, dies bei Zuzügen dann auch möglich ist.

Die verbleibenden Fehlerquellen sind falsche, nicht lokalisierbare Adressen und Adressen ohne Zielperson (z.B. leere Wohnungen), dieser Anteil beträgt in Österreich aber laut Tabelle 3.2 nur rund 4% der Adressen, dieser sehr niedrige Anteil bestätigt die hohe Qualität des Adressenmaterials von österreichischen Meinungsforschungsinstituten.

Tabelle 3.2: Qualität des Adressenmaterials für österreichische Umfragen<sup>47</sup>

Studie	I: Falsche Adressen	II: Adressen ohne Zielperson	I+II
Optima 1987	1,8%	0,9%	2,7%
Optima 1988	2,3%	1,1%	3,4%
Optima 1989	2,1%	1,1%	3,2%
Optima 1990	2,5%	1,5%	4,0%

---

<sup>46</sup> Vgl. [Költringer 1993] S. 112

<sup>47</sup> Vgl. Optima-Ausfallstatistik des Fessel-Instituts in [Költringer 1993] S. 114

Die weit gravierenderen Fehlerquellen bei Stichproben sind die Ausfallsfehler (Nonresponsefehler). Damit ist die Tatsache gemeint, daß nicht 100% der Personen der Ausgangsstichprobe befragt werden können. Bei vollständiger Befragung aller ausgewählten Stichprobenelemente spricht man von einer 100%igen Ausschöpfung. Normalerweise kann aber nur mit Ausschöpfungsraten bis zu 85% gerechnet werden.

Der Grad der Verzerrung durch Nonresponse steigt mit sinkender Ausschöpfungsraten und mit steigenden Merkmalsdifferenzen zwischen Nonrespondenten und Respondenten der interessierenden Statistik.

Drei Faktoren bestimmen den Nonresponse:

- Nichterreichbarkeit: Bestimmte mobile Bevölkerungsgruppen, z.B. junge, berufstätige Männer sind wesentlich schwieriger zu Hause anzutreffen als weniger mobiler, wie etwa Hausfrauen oder Pensionisten.
- Totalverweigerung oder Teilverweigerung: Manche Personen wollen sich überhaupt nicht interviewen lassen, andere lassen manche Fragen nicht zu und verweigern.
- Interviewunfähigkeit: Diese Personen werden oftmals aus der Zielpopulation ausgeschlossen, sie rechnet man zu den stichprobenneutralen Ausfällen. Es verbleiben also die beiden vorher erwähnten Nonresponse-Fehler übrig.

Individuelle Ausschöpfungsraten sind das wichtigste Kriterium zur Beurteilung von Interviewerleistungen. Die Anreizsysteme zur Erreichung hoher Ausschöpfungsraten darf man allerdings nicht zu stark einsetzen, da sonst der Anreiz zur Manipulation stärker wird.

Im internationalen Vergleich liegt Österreich bei den Ausschöpfungsraten relativ gut, wie aus Tabelle 3.3 hervorgeht:

## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

---

Tabelle 3.3: Ausschöpfungsraten nationaler Großumfragen (persönliche Befragungen) in Europa und Nordamerika <sup>48</sup>

Nation	Interviewdauer	Ausschöpfungsrate	Kontaktversuche
England			
Omnibus (1990)	35 min.	50%	5
NRS (1990)	37 min.	65%	5
Deutschland			
AGMA Press (1990)	50 min.	83%	8
AGMA Radio/TV (1990)	-	83%	8
Österreich			
Media-Analyse (1988)	50 min.	83%	8
Optima (1990)	50 min.	72%	-
USA			
MRI (1990)	50 min.	71%	6
SMRB (1990)	48 min.	72%	6

Den Nichterfassungsfehlern und Nonresponsefehlern wird mittels **Gewichtung** entgegengetreten.

Aufgrund disproportionaler Ausfälle in einigen Teilbereichen der Stichprobe ist für manche Merkmale eine Über-, und für manche eine Unterbesetzung zu erwarten. Das Sample wird nachträglich „gewichtet“, das heißt den bekannten Merkmalsverteilungen der Grundgesamtheit angeglichen. Sind z.B. nur 44% der Interviews von Männern, diese machen aber 47% der Grundgesamtheit aus, so müssen ihre Ergebnisse aufgewichtet werden.

Um das Ergebnis nicht zu stark zu beeinflussen, „*ist bei der Gewichtung der Faktor 1,5 als tolerable Grenze anzusehen.*“ <sup>49</sup>

---

<sup>48</sup> Vgl. [Költringer 1993] S. 123

<sup>49</sup> Vgl. [Karmasin/Karmasin 1977] S. 255

### ***III.3 Das Interview in der Umfrageforschung***

Das Interview ist die in der Soziologie am häufigsten verwendete Methode und hat einen breiten Anwendungsbereich. Es wird zur Basis der Auskünfte über Einstellungen, Meinungen und Verhalten der Befragungspersonen benützt. Es erscheint als einfache Methode, nicht zuletzt aufgrund seiner Nähe zum Alltagsgespräch. Im Detail betrachtet, gibt es aber zahlreiche Probleme und Designmöglichkeiten von Interviews.

Das Interview setzt gute verbale Fähigkeiten des Befragten und ein annähernd ähnliches Sprachspiel bei Forscher, Interviewer und Befragten voraus. In den meisten Fällen gibt es keinen persönlichen Kontakt von Forscher und Interviewer, so daß der Interviewer die eigentliche Kontaktperson des Befragten wird. Die Punkte, die der Forscher erheben will, stellt er kodiert als Fragen an die Zielperson. Er kodiert sie, indem er den Code, also die Umgangssprache und das Begriffsvokabular der Zielperson benützt.

Es entsteht eine Kette von Vermittlungen, Forscher - Fragebogen - Interviewer – Befragter, diese Kette kann wesentlich zur Verzerrung von Ergebnissen beitragen. Die zwei Stufen zwischen Forscher und Befragten – Fragebogen und Interviewer – kann man auch als Instrumente ansehen, mit der die Sachverhalte ermittelt werden sollen. Anders als in der Naturwissenschaft, haben diese Instrumente aber ein Eigenleben, und sind keine bloßen Meßwerkzeuge.

Das Hauptaugenmerk dieses Kapitels soll auf den Problemen im Bereich des Fragebogens (Frageformulierungen) und in der Interviewer-Befragten-Beziehung liegen.



### **III.3.1 Die Frage**

Fragen sind das gebräuchlichste Mittel, mit dem wir im Alltag versuchen, etwas zu ermitteln. Fragen werden aus einem Interesse des Forschers gestellt, die Antworten sollen Indikatoren der Variablen sein. *„Entscheidend ist, daß die Frage nur ein Mittel zur Antwort ist, diese erst die gesuchte Variable“*<sup>50</sup>

Die Formulierung der Fragen spielt eine wesentliche Rolle bei der Qualität einer Studie. Unter Formulierung versteht man:

- Frageneinleitung und Frage (Stimulus)
- Fragenstrategie und Antwortskala
- Schwierigkeit, Verständlichkeit und Präzision des Textes
- Positionierung der Frage im Fragebogen

### **III.3.2 Meßqualitätskriterien: Reliabilität und Validität**<sup>51</sup>

Unter Umfrage-Meßmethoden versteht man alle Datenerhebungs-, Formulierungs- und Positioniermerkmale eines Meßinstruments. Ein Meßinstrument ist in den folgenden Ausführungen auf einzelne Umfrage-Items bzw. Einzelfragen beschränkt. Zur Bewertung von Meßinstrumenten werden Meßqualitätskriterien herangezogen.

Um deren Bedeutung zu verstehen, muß zunächst etwas weiter ausgeholt werden. Bei einer Befragung können, grob gesprochen, zwei Arten von Fehler auftreten, nämlich Stichprobenfehler und Meßfehler. Unter Stichprobenfehler versteht man die in III.2 besprochenen, nämlich Statistische Zufallsfehler, Nichterfassungsfehler und Ausfallsfehler.

Meßfehler sind Interviewerfehler, Befragtenfehler und Meßmethodenfehler. Die ersten beiden werden später besprochen, dieser Abschnitt beschäftigt sich mit den

---

<sup>50</sup> Vgl. [Friedrichs 1980] S. 209

<sup>51</sup> Vgl. [Költringer 1993] S. 53-60

Meßmethodenfehlern. Diese beziehen sich sowohl auf die Datenerhebungsmethode als auch auf die Aspekte der Fragenformulierung.

Wie bei allen Fehlerarten, unterscheidet man variable Fehler und Verzerrungen (Bias). Wir nehmen an, man führe unabhängige Meßwiederholungen eines Umfragedesigns bei denselben Befragten durch, die Befragungen seien also replizierbar. Variable Meßfehler liegen vor, wenn dies zu unterschiedlichen Ergebnissen führt. Bias ist demgegenüber ein Fehlertyp, der die Ergebnisse aller Befragungen beeinflusst, ist also sozusagen ein konstanter Fehler.

Die Güte einer Meßmethode bzw. der Fragenformulierungen wird durch die variablen Fehler bestimmt. Zur Messung dieser zieht man das sogenannte True-Score-Meßmodell heran. Dieses relativ komplizierte Modell soll hier an einem einfachen Beispiel erläutert werden.

In der Umfrageforschung werden Fragen gestellt, um Antworten zu erhalten, die als Indikator für latente Variablen aufgefaßt werden. Oftmals interessieren nicht die „wahren“ Antworten (True-Score  $T_{ij}$ ), sondern die latenten Variablen  $F_i$ . Z.B. interessiert bei einer Recall-Frage („Welche Partei haben sie bei der letzten Wahl gewählt?“) nicht die tatsächliche Antwort, sondern die latente Variable „Welche Partei wird der Befragte bei der kommenden Wahl wählen?“, die den True-Score beeinflusst.

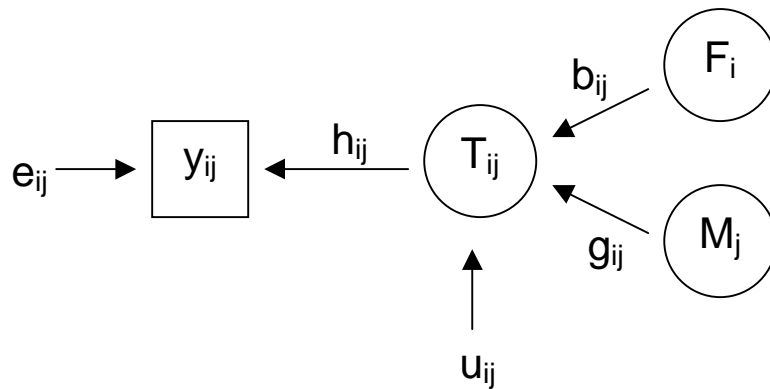
Weiters Einfluß auf den True-Score haben:

- Verwendete Meßmethode  $M_j$  (telefonisch oder persönlich, Art der Fragestellung, Zeigen einer Antwortkarte,...)
- Itemspezifische Effekte  $u_{ij}$ : Diese beschreiben alle anderen zeitlich stabilen Reaktionen, die nichts mit der zu messenden Variable zu tun haben. Z.B. wäre das in unserem Fall die Einstellung der Person zur Politik, zur Regierung,...

Die Antwort eines Befragten muß aber nicht mit dem True-Score übereinstimmen. Es können zufällige Fehler  $e_{ij}$  auftreten wie Mißverständnisse, Tipp- oder Verkodungsfehler, fehlerhafte Interpretationen der Fragen und vieles mehr.

Der beschriebene Zusammenhang ist in Abbildung 3.5 dargestellt.

Abbildung 3.5: Darstellung des True-Score-Meßmodells



$y_{ij}$  ..... Antwortreaktion auf das Meßinstrument für Frage  $i$  unter der Verwendung der Meßmethode  $j$  (standardisiert:  $E(y_{ij})=0$ ,  $\sigma(y_{ij})=1$ ),

$T_{ij}$  ..... True-Score für das Konstrukt  $i$  und die Meßmethode  $j$ ,

$F_i$  ..... Latente Variable  $i$  (standardisiert:  $E(F_i)=0$ ,  $\sigma(F_i)=1$ ),

$M_j$  ..... Meßmethodenfaktor für die Meßmethode  $j$  (standardisiert:  $E(M_j)=0$ ,  $\sigma(M_j)=1$ ),

$e_{ij}$  ..... Zufälliger Meßfehler ( $E(e_{ij})=0$ ,  $\sigma(e_{ij})$ ),

$u_{ij}$  ..... Itemspezifischer Faktor ( $E(u_{ij})=0$ ,  $\sigma(u_{ij})$ ),

$h_{ij}$  ..... Korrelation zwischen  $T$  auf  $y$

$b_{ij}$  ..... Korrelation zwischen  $F$  auf  $T$

$g_{ij}$  ..... Korrelation zwischen  $M$  auf  $T$

Die verschiedensten Einflüsse müssen nun noch mathematisch formuliert werden, um gemessen werden zu können. Es wird angenommen, daß die Faktoren  $F$ ,  $M$  und  $u$  untereinander und vom Zufallsfehler  $e$  unabhängig sind. Dies führt dazu, daß für alle  $i, j$  gilt:  $\text{Cov}(F_i, M_j) = \text{Cov}(F_i, u_{ij}) = \text{Cov}(M_j, u_{ij}) = \text{Cov}(F_i, e_{ij}) = \text{Cov}(M_j, e_{ij}) = \text{Cov}(u_{ij}, e_{ij}) = 0$ .

Will man verschiedene Designs prüfen, so muß man eine Varianzanalyse durchführen, auf die hier nicht näher eingegangen werden kann.

Wie erläutert, wird beim Zusammenhang zwischen  $y_{ij}$  und  $T_{ij}$  davon ausgegangen, daß nur zufällige Meßfehler vorliegen. Die gleichen Fragen werden von denselben Personen unter exakt gleichen Bedingungen mehrmals beantwortet, die

Meßwiederholungen seien voneinander unabhängig und es kommt zu keinen Veränderungen (z.B. Einstellungsänderungen) zwischen den Meßwiederholungen.

Differenzen können nur auf zufällige, zeitlich instabile Faktoren zurückgeführt werden, das heißt:  $y_{ij} = T_{ij} + e_{ij}$  mit  $\text{Cov}(T_{ij}, e_{ij}) = 0$  und  $E(e_{ij}) = 0$ .

Aufgrund der Nullkorrelation zwischen True-Score und zufälligen Meßfehlern ergibt sich folgende Varianzzerlegung:  $\sigma(y_{ij}) = \sigma(T_{ij}) + \sigma(e_{ij})$

Die Test-Retest-Reliabilität definiert sich dann mit:

$$\text{Test-Retest-Reliabilität} = \sigma(T_{ij}) / \sigma(y_{ij})$$

Ergibt eine Varianzanalyse eines Befragungsdesigns eine Reliabilität von 0,8, dann sind 80% der Varianz der Messungen zeitlich stabil, und 20% beruhen auf zufälligen Meßfehlern. Mit der Korrelation zwischen  $y_{ij}$  und  $T_{ij}$ ,  $h_{ij} = r(y_{ij}, T_{ij})$ , ergibt sich:

$$\text{Test-Retest-Reliabilität} = \sigma(T_{ij}) / \sigma(y_{ij}) = h_{ij}^2$$

Der Begriff Reliabilität bezieht sich also auf den standardisierten, linearen Zusammenhang zweier Meßergebnisse für dieselbe Frage, unabhängig von True-Score-Veränderungen zwischen den Meßwiederholungen. Eine hohe Reliabilität bedeutet, daß der Einfluß von zufälligen Fehlern wie Mißverständnissen gering ist.

Invalide Meßmethodeneffekte beruhen darauf, daß die Befragten zeitlich stabil auf die verwendete Meßmethode, z.B. auf eine vorgegebene Antwortkarte, reagieren, unabhängig von ihren Werten in der zu messenden Variablen.

Invalide itemspezifische Effekte beschreiben alle anderen zeitlich stabilen Reaktionen, die nichts mit der zu messenden latenten Variable zu tun haben.

Das erweiterte True-Score-Modell ist:  $y_{ij} = h_{ij} * T_{ij} + e_{ij}$  mit  $T_{ij} = F_j * b_{ij} + M_i * g_{ij} + u_{ij}$ .

Die Varianzzerlegung für den standardisierten True-Score ist:

$$\sigma(T_{ij}) = \sigma(u_{ij}) + b_{ij} + g_{ij} = 1$$

Die Komponente  $b_{ij}^2$ , die **True-Score-Validität**, gibt jenen Varianzanteil des True-Scores an, der auf die gemessene Variable zurückzuführen ist. Gute Designs haben hohe  $b_{ij}$ -Werte und niedrige  $\sigma(u_{ij})$ - und  $g_{ij}$ -Werte, das heißt, daß der Einfluß von Meßmethode und itemspezifische Effekte auf den True-Score gering sind. Ergibt eine Varianzanalyse, daß  $b_{ij}^2=0,8$ ,  $\sigma(u_{ij}) = g_{ij} = 0,1$ , so bedeutet dies, daß 80% der True-Score-Varianz zeitlich stabil sind, und daß je 10% auf Einflüsse der verwendeten Meßmethode und auf itemspezifische Effekte zurückzuführen ist.

Die Komponente  $g_{ij}^2$ , die **True-Score-Meßmethodenvarianz**, gibt jenen Varianzanteil des True-Scores an, der durch die konkrete Meßprozedur  $M_i$  erklärt wird. In der Literatur findet man auch oft die Begriffe *Indikator-Validität* (*Konstruktvalidität*,  $= h_{ij}^2 * b_{ij}^2$ ) und *Indikator-Meßmethodenvarianz* ( $= h_{ij}^2 * g_{ij}^2$ ).

Die abgeleiteten Kennzahlen sind nützliche und in der Praxis auch angewandte Kriterien zur Beurteilung der Meßqualität. Man wird jenen Meßinstrumenten den Vorzug geben, die sich durch geringe invalide Varianzkomponenten, hohe Validität und Reliabilität auszeichnen, das heißt, daß latente Variablen großen Einfluß auf den True-Score haben und daß zufällige Fehler (Mißverständnisse, Tippfehler) nur geringen Einfluß auf die Antwortreaktion haben.

Insgesamt ist das Meßqualitätsniveau in Österreich sehr hoch, wie aus Tabelle 3.4 im Vergleich mit der USA ersichtlich ist. Es läßt sich daraus schließen, daß die Umfrageprozeduren führender Umfrageinstitut einem hohen Standard entsprechen. Positiv dürfte sich auch die sprachlich-kulturelle Homogenität der österreichischen Bevölkerung auswirken.

Tabelle 3.4: Vergleich des Meßqualitätsniveaus zwischen Österreich und den USA<sup>52</sup>

Meßqualitätskriterium (Mittelwerte)	Österreich	USA
Indikator-Validitätseffekt	0,85	0,81
Indikator-Meßmethodeneffekt	0,2	0,26
Effekt zufälliger Meßfehler	0,49	0,52

---

<sup>52</sup> Vgl. [Költringer 1997] S. 59

Das Quadrat dieser Kennwerte entspricht den jeweiligen Varianzanteilen in den Antwortverteilungen. In Österreich sind 72% ( $0,85^2$ ) der Varianz valide, in den USA 66% ( $0,81^2$ ).

Welche Aspekte des Umfragedesigns die Meßqualität beeinflussen wird im folgenden behandelt.

### ***III.3.3 Formulierung von Fragen***

#### **a) Einleitung**

Mit der Einleitung meint man jenen Teil der Befragung, in der sich der Interviewer vorstellt und den Befragten zur Teilnahme am Interview bittet. Es handelt sich um eine sehr wichtige Phase im Gespräch, hier wird die Beziehung zwischen Interviewer und Befragten installiert, außerdem wird der generelle Bezugsrahmen des Befragten durch die Beschreibung des Untersuchungszieles und des Auftraggebers festgelegt. All dies kann das Verhalten des Befragten während des gesamten Interviews sehr beeinflussen.

Dieser Teil des Interviews läßt sich in der Praxis nicht standardisieren, die Interviewer sind gezwungen, ihn frei zu handhaben, da die Rückfragen äußerst unterschiedlich sein können, und ihre Beantwortung notwendig zur Durchführung des Interviews sein können.

Standardisieren läßt sich allerdings der schriftliche Vorspann, und der sollte enthalten:<sup>53</sup>

- den Namen und nötigenfalls eine kurze Beschreibung des Instituts,
- den Zweck, das Ziel und den Auftraggeber der Untersuchung,
- die Zusicherung der Anonymität und
- eine kurze Erläuterung, warum der Befragte ausgewählt wurde.

---

<sup>53</sup> Vgl. [Karmasin/Karmasin 1977] S. 196

Der Zweck der Untersuchung kann als „wissenschaftlich“ angegeben werden, dadurch wird meist ein positives Motiv zur Teilnahme erreicht.

Der Auftraggeber sollte allerdings nicht explizit genannt werden, es ist sonst mit einer positiven Verzerrung der Ergebnisse zugunsten des Auftraggebers zu rechnen.

Das Ziel der Untersuchung sollte auch nicht zu genau spezifiziert werden, da der Befragte sonst versucht, die im Hinblick auf dieses Ziel „richtigen“ Antworten zu liefern. Bei zu vager Formulierung des Zieles versucht der Befragte, das Ziel während des Gesprächs „herauszufinden“, dies kann seine Aufmerksamkeit stark in Mitleidenschaft ziehen.

Die Einleitung sollte wie das ganze Interview locker und zwanglos verlaufen. Der Eindruck einer echten Gesprächssituation sollte erzeugt werden, die der Befragte interessant findet.

#### b) Fragen, Fragestrategien und Antwortvorgaben

Fragen können offen oder geschlossen gestellt werden. Darunter versteht man, ob eine oder mehrere Antworten vorgegeben werden (geschlossen), oder eben nicht (offen).

Obwohl die offene Frage auch einige Vorteile aufweist, hat sich die geschlossene Frage bei den meisten Anwendungen aufgrund ihrer leichten Auswertbarkeit und Vergleichbarkeit der Antworten durchgesetzt. Bei einer offenen Frage, bei der es dem Befragten natürlich möglich ist, viel genauer und präziser seine Einstellungen kundzutun, müssen die heterogenen Antwortklassen erst in Kategorien eingeteilt, codiert werden, dies ist wesentlich teurer.

Eine weiteres Argument für geschlossene Fragen ist, „...*daß sehr viele Menschen von sich aus nicht in der Lage sind, ihre Ansicht zu bestimmten Problemen zu formulieren und der Artikulationshilfe bedürfen.*“<sup>54</sup>

---

<sup>54</sup> Vgl. [Holm 1982] S. 99

Bei den Antwortvorgaben unterscheidet man zwischen internen und externen. Unter intern versteht man z.B.: „Ist Oberösterreich ein Land mit einem hohen, mittleren oder niedrigen Wohlstand?“<sup>55</sup>

Würde die Frage lauten: „Wie schätzen sie den Wohlstand in Oberösterreich ein? a)hoher Wohlstand, b)mittlerer Wohlstand, c)niedriger Wohlstand“, dann spräche man von externer Antwortvorgabe. Die interne sollte nur verwendet werden, wenn sie aus zwei, drei leicht merkbaren Alternativen besteht.

Bei der Formulierung von Antwortkategorien ist darauf zu achten, daß die Abstufungen richtig gewählt werden. Es sind die gleichen Regeln wie bei der Klassifikation von Merkmalen heranzuziehen, um die Ausprägungen der Variablen korrekt zu messen. Wichtige Einflüsse auf die Qualität des Meßergebnisses haben:<sup>56</sup>

- 1) Länge der Antwortskala ( $\beta^{57}=0,64$  bei der Reliabilität und  $\beta=0,51$  bei der Validität): Sie hat den stärksten Effekt auf die Meßqualität der Daten. Während zwei- bis dreistufige Skalen wegen ihrer schlechten Meßeigenschaften kaum attraktiv sind, weisen sehr lange kategorielle Skalen (neun- bis elfstufig) hohe invalide Meßmethodeneffekte auf. Vier- bis siebenstufige Instrumente stellen die beste Wahl dar.
- 2) Präsentation von Antwortkarten: Éntgegen gängigen Annahmen haben Antwortkarten keinen bedeutenden Einfluß auf Reliabilität und Validät. Sie reduzieren aber in beträchtlichem Maße den relativen Bias des Mittelwerts.
- 3) Symmetrische Antwortskalen: Sie reduzieren klarerweise den relativen Mittelwertsbias im Vergleich zu asymmetrischen Skalen, haben aber keinen Einfluß auf die Zuverlässigkeit und Gültigkeit der Messungen.

---

<sup>55</sup> Vgl. [Spectra-Studie 9/97] S. 1

<sup>56</sup> Vgl. [Költringer 1997] S. 107-109

<sup>57</sup>  $\beta$  mißt die Stärke des Zusammenhangs zwischen Reliabilität bzw. Validität und Umfragedesign



- 4) Explizierte Mittelkategorie: Der Befragte muß sich nicht für ein „Extrem“ entscheiden. Die Vorgabe einer Mittelkategorie hat keinen Einfluß auf Reliabilität und Validität, reduziert aber den relativen Bias des Mittelwerts.
- 5) Explizierte „Weiß nicht“ oder „keine Angabe“-Filter: Dadurch hat der Befragte, der keine Meinung oder Information zu einem Befragungsthema hat, die Möglichkeit, nicht zu antworten. Andererseits können es sich Befragte dadurch natürlich leichtmachen. Interviewer sollten angewiesen werden, bei einer „Weiß nicht“-Antwort nachzufragen und auf eine „gültige“ Antwort drängen. Dann reduziert sich der hohe „Weiß-nicht“-Anteil, und die Vorgabe hat laut amerikanischen Studien einen positiven Effekt auf die Datenqualität.

Keinen Einfluß auf die Meßqualität haben verbale Bezeichnungen von Antwortkategorien oder die Richtung der erstgenannten Skalenkategorie.

Eine typisches Befragungssitem besteht, abgesehen von der Antwortskala, aus einer kurzen Einleitung und einem Stimulus (Fragesatz), auf den die Befragten reagieren sollen. Fragen sollten kurz, einfach und auf den Bezugsrahmen des Befragten bezogen sein. In diesem Zusammenhang wird die Qualität der Daten beeinflusst durch.<sup>58</sup>

- 1) Länge von Frageneinleitungen und Stimulus ( $\beta=0,37$  bzw.  $\beta=0,38$ ): Die Reliabilität von Messungen wird durch kurze Frageneinleitungen (bis 24 Wörter) und kurze Stimuli (4-13 Wörter) optimiert. Die Validität der Messungen wird durch mittellange Einleitungen (25-54 Wörter) und durch lange Fragen (14 und mehr Wörter) maximiert. Durch diese gegenläufigen Einflüsse ist es schwer, ein klare Aussage zu treffen. Da kurze Einleitungen und Stimuli allerdings auch den kleinsten relativen Mittelwertsbias aufweisen, stellen sie wohl die beste Meßmethodenwahl dar.

---

<sup>58</sup> Vgl. [Költringer 1997] S. 109-111

- 2) Ausgewogenheit der Formulierung: Bei internen Antwortvorgaben können sich Anzahl und Intensität die Waage halten („Wie zufrieden oder unzufrieden sind sie mit der politischen Situation“), oder es liegt ein unausgewogener Stimulus vor („Wie zufrieden sind sie mit der politischen Situation“). Die Ausgewogenheit hat keinen Effekt auf Meßqualität und den relativen Bias des Mittelwerts.
- 3) Absolute versus vergleichende Formulierung: Amerikanische Studien zeigen, daß mit vergleichenden Fragenformulierungen bessere Daten erzielt werden. Österreichische Studien können dies aber für heimische Umfragen nicht bestätigen.

Um Fragen verständlich zu gestalten sollte weiters auf kurze Worte, auf Worte der Muttersprache, Vermeidung von Fremdwörtern, Verwendung von Konkreta gegenüber Abstrakta, persönliche gegenüber „logischen“ Formulierungen und Sätze mit geringem Komplexitätsgrad geachtet werden.

### c) Aufbau des Fragebogens und Positionierung einer Frage

Die ersten Fragen sollten eher neutral sein und das Interesse des Befragten wecken bzw. verstärken. Schwierige oder mit möglichen Verweigerungen belastete Fragen sollte man eher in der Mitte oder am Ende stellen.

Der sequentielle Aufbau des Fragebogens ist von großer Bedeutung, der Befragte sollte einen Zusammenhang der Fragen konstruieren können. Es gibt eine Reihe von praktischen Erfahrungen der Institute über den „richtigen“ Aufbau einer Befragung. Hier soll etwas näher auf den Einfluß der Position einer Frage im Fragebogen und auf die Verwendung von „Fragenbatterien“<sup>59</sup> eingegangen werden:<sup>60</sup>

---

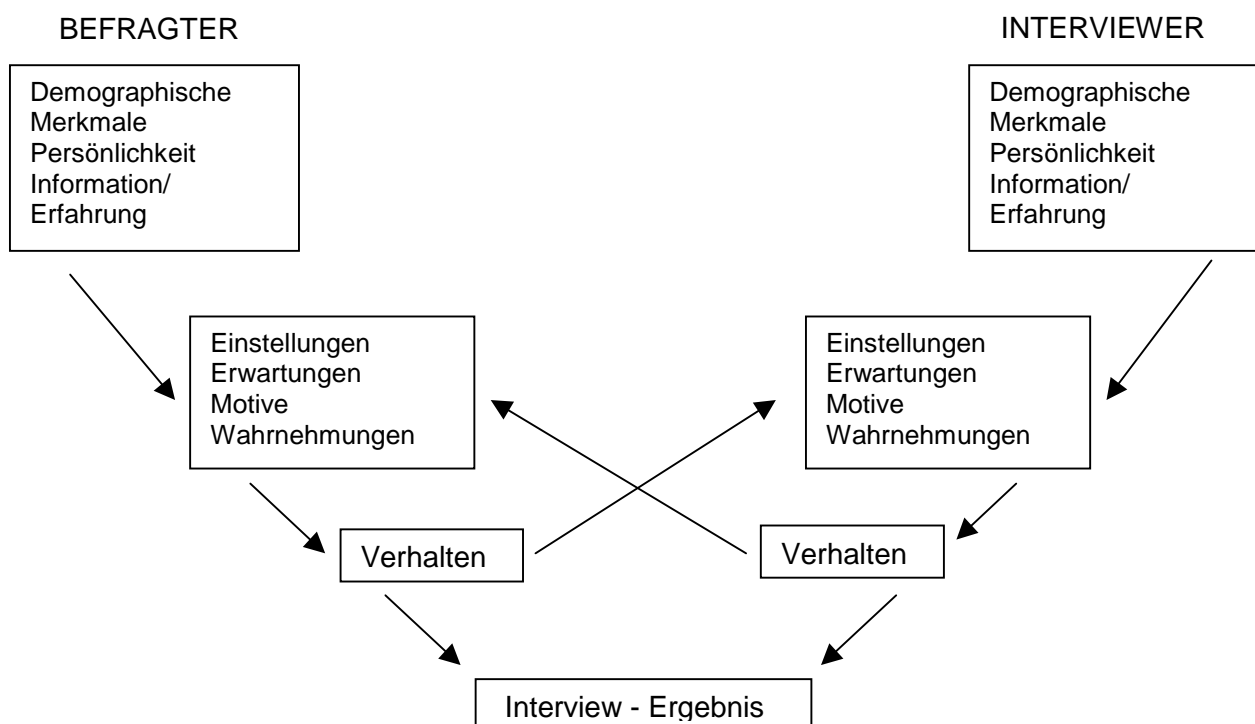
<sup>59</sup> Unter einer Fragenbatterie versteht man das Abfragen von mehreren Fragen in einer Formulierung. Z.B.: „Welche Partei hat ihrer Meinung nach am meisten für .... getan? Sicherung der Arbeitsplätze, Bekämpfung der Kriminalität,...“

<sup>60</sup> Vgl. [Költringer 1997] S. 111,112

- 1) Länge der Fragenbatterie ( $\beta=0,36$  bei Reliabilität): Diese Umfragestrategie erlaubt ein schnelles Abfragen zahlreicher Stimuli, wird also wegen ökonomischen Überlegungen angewandt. Diesen Vorteil erkaufte man sich aber durch Verlust der Datenqualität, da der Charakter von Fragebatterien zu einer erhöhten Nachlässigkeit von Interviewer und Befragten führen dürfte. Batterien weisen deutlich niedriger Reliabilitäten auf als Einzelfragen, dieser negative Effekt liegt aber bei der Validität der Messungen und dem relativen Mittelwertsbias nicht vor.
- 2) Position der Frage im Fragebogen ( $\beta=0,36$  bzw.  $\beta=0,47$ ): Die schlechtesten Daten erhält man am Beginn und am Ende eines längeren (>1Stunde) Interviews. Reliabilität und Validität sind hier am geringsten.  
Am Beginn dauert es eine gewisse Zeitspanne bis eine produktive Interviewersituation installiert ist, am Ende läßt die Konzentration der Befragten zu wünschen übrig.

### III.3.4 Interviewer-Befragten-Beziehung

Abbildung 3.6: Das Interview als sozialer Prozeß <sup>61</sup>



Interviewer tragen in vielfältiger Weise zum Befragungsergebnisse bei. Sie kontaktieren die Zielperson, motivieren diese zum Interview, sie lesen die Fragen vor, beantworten Rückfragen,....

Obwohl sie während der Befragung kontinuierlich mit dem Befragten interagieren, müssen sie jede Beeinflussung des Interviewpartners vermeiden. Die Steuerung des Kommunikationsprozesses ist von großer Bedeutung. Um einheitliche, standardisierte Interviews zu erreichen, gibt es eine Vielzahl von Verhaltensregeln für den Interviewer.

Der Interviewer-Einschulung kommt zentrale Bedeutung zu. Hierbei gibt es sehr unterschiedliche Strategien einzelner Institute. Während größere, namhafte Institute Wert auf gute Einschulung und einen langen Verbleib des Interviewers legen, setzen kleine Institute häufig auf die „Hire-and-Fire“-Strategie. Sie haben einfach nicht die finanziellen Möglichkeiten, häufig werden Interviewer dann auch nur schriftlich eingeschult, es erfolgt kein direkter persönlicher Kontakt.

Die Intensität der Interviewerschulung und der Einsatz erfahrener Interviewer hat einen positiven Effekt auf die Qualität der Daten, auf Validität und Reliabilität. Der Einfluß ist aber nicht so stark wie der von Fragebogendesign und Fragenformulierungen. Die vier Ursachen für interviewerabhängige Meßfehler sind:<sup>62</sup>

- 1) Absichtliches Fehlverhalten: Dies kann nie völlig ausgeschlossen und unterstreicht die Wichtigkeit von Einschulung und regelmäßigen Kontrollen, diese werden telefonisch, postalisch oder persönlich vorgenommen.
- 2) Äußere Merkmale des Interviewers: Bei der Interviewer-Befragten-Interaktion wirken die gleichen Gesetzmäßigkeiten wie bei allen anderen sozialen Interaktionen. Die sichtbaren demographischen Merkmale des Interviewers können das Befragtenverhalten beeinflussen.

---

<sup>61</sup> Vgl. [Friedrichs 1980] S. 218

- 3) Bearbeitung des Fragebogens: Eine Variabilität der Befragungssituation entsteht, wenn verschiedene Interviewer Fragen nicht vorlagengetreu präsentieren, verschiedene Satzteile unterschiedlich betonen,...
- 4) Assistenzleistung des Interviewers: Wenn der Befragte eine Frage nicht genau versteht und nachfragt, so gibt ihm der Interviewer Hilfestellung. Diese ist nicht genau zu standardisieren, die dann folgenden Antwortreaktionen können wieder interviewerabhängig sein.

Den Einfluß des Interviewers erhält man folgendermaßen:<sup>63</sup> Zunächst wird der Meßwert als Summe von dem wahren Wert einer Person, dem Zufallsfehler und dem Effekt eines Interview auf die Reaktion auf eine Person definiert.

$$y_{ij} = X_i + M_{ij} + e_{ij}$$

$y_{ij}$  ....Meßwert der i-ten Person und des j-ten Interviews

$X_i$  ....Wahrer Wert der i-ten Person

$M_{ij}$  ...Effekt des j-ten Interviews auf die Reaktion der i-ten Person

$e_{ij}$  ....Zufallsfehler der i-ten Person und des j-ten Interviews

$E(M_{ij})=E(e_{ij})=0$  und  $Cov(X_i, e_i)= Cov(M_i, e_i)= Cov(X_i, M_{ij})=0$ .

$$\rho = \sigma(M_{ij}) / \sigma(y)$$

$\rho$  gibt also den relativen Anteil der Interviewervarianz im Verhältnis zur Gesamtvarianz der Meßwerte an.  $\rho$  beträgt im Durchschnitt etwa 2% bei persönlichen Umfragen, und 1% bei Telefoninterviews.

Der zweite Part beim Interview, der Befragte, hat natürlich auch Einfluß auf die Qualität der Daten. Statistisch signifikante Unterschiede bezüglich der Reliabilität und der Validität gibt es bei der Befragung älterer Personen und bei der Befragung von Personen mit niedriger Bildung. Keine Unterschiede sind bei Personen mit

---

<sup>62</sup> Vgl. [Költringer 1992] S. 6

<sup>63</sup> Vgl. [Költringer 1992] S. 7

Befragungserfahrung und bestimmten Einstellungen zur Umfrageforschung zu bemerken.

Diese Einflüsse der Befragten sind insgesamt äußerst gering. *„Schlechte Umfragedaten werden überwiegend durch schlechte Forschungsdesigns verursacht, nicht durch die Unfähigkeit oder Unwilligkeit der Befragten“*<sup>64</sup>

### **III.3.5 Das Telefoninterview**

Das Telefoninterview stellt eine oft angewendete Sonderform bei der Befragung von Personen dar. Dabei wird die Zielperson vom Interviewer angerufen und gebeten, am Telefon zu gewissen Themen Stellung zu nehmen.

Das Interview sollte nicht länger als 10-15 Minuten sein, um den Charakter eines „normalen“ Telefongesprächs zu wahren. Dadurch beschränkt sich der Inhalt auf schnell abfragbare Themen, bei denen kein langes Nachdenken und sorgfältiges Abwägen erforderlich ist.

Der Vorteil liegt bei der Kosten- und Zeitersparnis und im Gegensatz zu schriftlichen Befragungen ist bei Telefoninterviews die Bereitwilligkeit, sich befragen zu lassen, relativ hoch, da ein Stimmkontakt in höherem Ausmaß als persönliches Ansprechen erlebt wird.

Allerdings erreicht man dadurch nicht alle Bevölkerungsgruppen, vor allem sozial schwache, sozial bessergestellte (Geheimnummern) und ländliche Bevölkerungsgruppen sind in Telefonumfragen unterrepräsentiert. Dies läßt sich aber durch Gewichtung in den meisten Fällen ohne Probleme ausgleichen.

Ein Effekt der geringeren persönlichen Bindung von Interviewer und Befragten bei Telefoninterviews ist, daß heikle Fragen gerne verzerrt beantwortet werden.

---

<sup>64</sup> Vgl. [Költringer 1993] S. 89

### **III.4 Umfragen zur Oberösterreichischen Landtagswahl**

#### **III.4.1 Umfragen zu politischen Wahlen**

Wie schon in Kapitel III.1 angesprochen, werden vor politischen Wahlen in gewissen Zeitabständen Meinungsumfragen erstellt, um einen gewissen Trend erkennen zu können. Das Erkennen eines Trends erfordert viel Erfahrung des Instituts, da die Veränderungen meist immer innerhalb der statistischen Schwankungsbreiten liegen. Z.B. war der Abwärtstrend der SPÖ in Oberösterreich für IFES schon erkennbar, obwohl dies nicht unbedingt aus den Zahlen herauszulesen war, wie wir später sehen werden.

Der Zweck einer Meinungsumfrage richtet sich naturgemäß nach den Zielen des Auftraggebers. Ist der Auftraggeber eine politische Partei, so ist der Zweck nicht, ein möglichst genaues Wahlergebnis zu prognostizieren, sondern um strategische Ausrichtungen treffen zu können und um ungefähre Trends zu erfahren.

Bei der strategischen Ausrichtung ist vor allem wichtig, mit welchen Personen man kandidiert, welche Eigenschaftsprofile man abdecken will, mit welchen Themen man in den Wahlkampf geht, welche Kompetenzen herausgestrichen werden sollen, welche Schwachpunkte der Gegner hat, soll kämpferisch oder zurückhaltend in den Wahlkampf gegangen werden,....

Das heißt, es werden gar keine allzu großen Anstrengungen unternommen, das Wahlergebnis möglichst genau vorherzusagen.

*„Es ist nicht unser Interesse, Wahlprognosen zu machen. Das ist ein journalistisches Interesse. Es ist auch kein Interesse des Auftraggebers.“<sup>65</sup>*

---

<sup>65</sup> Georg Michenthaler von IFES im Interview vom 28.10.97

Da es sich bei der Prognose des Wahlergebnisses also hauptsächlich um ein mediales Interesse handelt, wird diesem Punkt mehr Interesse zuteil, wenn Medien der Auftraggeber sind. Dies ist aber meist nur bei österreichweiten Umfragen der Fall, vor Nationalratswahlen z.B. geben Zeitschriften wie *News* oder *Profil* genaue Prognosen in Auftrag. Die Institute haben dann gewisse Möglichkeiten, das genaue Wahlergebnis vorherzusagen.

Das Hauptproblem bei der Erstellung von Wahlprognosen in der heutigen Zeit ist, daß die Anzahl der Nicht-Deklarierten bei Umfragen stark im steigen ist. Viele von ihnen werden trotzdem zur Wahl gehen, entscheiden sich aber erst im letzten Moment, sind also sogenannte Late-decider.

Z.B. hat eine Nachwahlenanalyse in einer großen niederösterreichischen Stadt 1997 ergeben, daß sich 9% am letzten Tag, weitere 20% in der letzten Woche vor der Wahl für die letztlich gewinnende Partei entschieden haben.

Was macht man also mit der großen Anzahl der Nicht-Deklarierten. IFES hat bei der OÖLTW einfach angenommen, die Nicht-Deklarierten verhalten sich so wie die Deklarierten. Dies ist natürlich eine Fiktion und führt nur zu sehr vagen Ergebnissen. Da es aber, wie schon erwähnt, nicht Interesse des Auftraggebers war, ein möglichst genaues Ergebnis zu prognostizieren, ist diese Fiktion ausreichend.

Um ein genaueres Ergebnis zu erhalten, versucht man eine Regressionsbeziehung aufzustellen. Rufen wir uns das True-Score-Modell wieder in Erinnerung. Die Antwort auf gewisse Fragen hängt von latenten Variablen ab. Man kann also die mögliche Wahlentscheidung einer Person durch das Abfragen von anderen Zusammenhängen „herauskitzeln“ indem man diese Parameter durch Korrelationskoeffizienten in Verbindung setzt. Dies erfordert viel Erfahrung eines Instituts und wird natürlich nicht verraten, da jedes Institut seinen eigenen Erfahrungswerte nötigenfalls einsetzt.

Ein Regressionsmodell könnte man sich vereinfacht aber so vorstellen: Es gibt gewisse Fragen, deren Antwort eng mit der Wahlentscheidung zusammenhängen. Etwa die Recall-Frage: „Welche Partei haben sie bei der letzten Landtagswahl



gewählt?“ Im einfachen Fall kann man das Ergebnis dieser Frage mit dem tatsächlichen Ergebnis, das man ja genau kennt, vergleichen und entsprechend das Ergebnis der „neuen Sonntagsfrage“ auf- oder abgewichten.

Geht man etwas genauer vor, so kann man aus Erfahrungswerten sagen, daß das Ergebnis der Recall-Frage und die Wahlentscheidung mit einem gewissen Korrelationskoeffizienten in Verbindung stehen. Nimmt man noch andere Parameter in das Modell hinein, man setzt etwa „Zufriedenheit mit der Lebenssituation in Oberösterreich“ mit dem Wählen einer Regierungspartei in Verbindung, so kann man Modelle erstellen, die das Zuteilen der Nicht-Deklarierten zu einer gewissen Partei ermöglichen.

Dazu kommt noch, daß man aus Erfahrung weiß, zu welchen politischen Parteien sich Personen sehr ungern deklarieren. Etwa hat es in der Vergangenheit eine Tendenz gegeben, daß sich FPÖ-Wähler sehr ungern deklarieren, dies ändert sich allerdings schon. In kleinen Gemeinden mit starker ÖVP-Dominanz deklarieren sich SPÖ-Wähler äußerst ungern, dies war auch in Oberösterreich zu beobachten.

Wie genau man eine Prognose durch das Heranziehen dieser Methoden stellen kann, läßt sich für die Oberösterreich-Wahl leider nicht nachvollziehen, da diese nicht herangezogen wurden.

### ***III.4.2 Ergebnisse der Institute bei Prognosen für die OÖLTW<sup>66</sup>***

Bei der Oberösterreichischen Landtagswahl haben das Institut für empirische Sozialforschung (IFES) für die SPÖ und das Marktforschungsinstitut SPECTRA für die ÖVP Meinungsumfragen durchgeführt, die sie für diese Arbeit zur Verfügung gestellt haben. Interessant ist, daß SPECTRA teilweise Analysen für die ÖVP und die *Oberösterreichischen Nachrichten* erstellt hat.

---

<sup>66</sup> Vgl. IFES-Studien 7/97 bis 9/97 und SPECTRA-Studien 7/97 bis 9/97

Bei den folgenden Betrachtungen werden die im Zeitraum von Ende Juni, Anfang Juli bis Ende September 1997 durchgeführten Befragungen herangezogen, also auch der Zeitraum in dem der APSM97 lief.

a) Strukturierung des Objektbereichs

Klassifikation und Nominalskalen bei der OÖ-LTW (Grundgesamtheit: Bevölkerung ab 18 Jahren) durch IFES:<sup>67</sup>

- A) Alter: - bis 29 Jahre .....1  
- 30 bis 49 Jahre .....2  
- 50 bis 69 Jahre .....3  
- 70 Jahre und älter .....4
- B) Geschlecht: - männlich .....1  
- weiblich .....2
- C) Abgeschlossene Schulbildung:  
- Pflichtschule ohne weitere Ausbildung .....1  
- Pflichtschule mit weiterer Ausbildung .....2  
- Matura, Hochschule .....3
- D) Tätigkeit:  
- Schüler/Student .....1  
- nicht berufstätig / im Haushalt .....2  
- Pension .....3  
- Selbständig oder freiberuflich tätig .....4  
- Arbeiter (+Lehrlinge für Arbeiterberufe) .....5  
- Angestellte (+Lehrlinge für Angestelltenberufe) ...6  
- Öffentlich Bedienstete .....7
- E) Region: - Linz und Umgebung .....1  
- Innviertel .....2  
- Mühlviertel .....3  
- Traunviertel .....4  
- Hausruckviertel .....5

Das SPECTRA-Institut hat in ihren Untersuchungen den Objektbereich wesentlich genauer strukturiert<sup>68</sup>. Mehr Klassen gab es für die Merkmale Region (die 18 politischen Bezirke von OÖ), Alter (Elf Klassen, von 18 bis 70 Jahre in etwa

---

<sup>67</sup> Vgl. [IFES-Studie 9/97] S. 4

<sup>68</sup> Vgl. [SPECTRA-Studie 9/97] Anhang S. 8-10

fünfjährigen Abständen und ab 70 Jahre) und Ausbildung (Fünf Klassen mit Pflichtschule, Lehrausbildung, Berufsbildende mittlere Schule, Matura und Hochschule).

Die Klassifizierung der Tätigkeit und des sozialen Status erfolgte bei SPECTRA über:

Sind sie berufstätig? - Ja .....1  
- Nein.....2  
- Keine Angabe....3

Welcher Berufsgruppe gehört der Hauptverdiener des Haushalts an, oder hat angehört (falls in Pension)?

- Arbeiter, Facharbeiter .....1
- Angestellter .....2
- Beamter .....3
- Selbständig .....4
- Landwirt .....5
- Andere Angabe .....6

Liegt das gesamte Netto-Monatseinkommen in Ihrem Haushalt (alle Verdiener zusammengerechnet) –

- unter 24.000,- .....1
- über 24.000,- .....2
- Keine Angabe.....3

Wieso gerade diese Klassifizierungen, und nicht andere gewählt wurden, ist Ergebnis einer Optimierung der Befragung in ökonomischer und qualitativer Hinsicht. In qualitativer Hinsicht haben die Institut ihre Erfahrungswerte, welche Klassifikation bessere Ergebnisse liefert.

Gewissen Einfluß hat natürlich auch die Stichprobengröße und die Befragungsart. Während IFES nur telefonische Befragungen mit Stichprobeneinheiten von 423-705 durchführte, hat SPECTRA telefonische Befragungen mit bis zu 1500 Interviews und persönliche Befragungen mit 500-1000 Befragungen durchgeführt.

Interessant ist, ob der erhebliche Mehraufwand des SPECTRA-Institut gegenüber IFES auch zu genaueren Prognosen geführt hat.

## b) Aufbau der Fragebögen und Frageformulierungen

Beide Institute haben die Fragen nach der Statistik am Ende des Fragebogens gestellt, wo man mit den geringsten Verzerrungen rechnen kann.

Während SPECTRA den Einstieg relativ sanft mit „Leben Sie gerne in Oberösterreich oder würden Sie lieber woanders wohnen“ gestaltete, kam IFES relativ schnell zur Sache: „Von welcher oberösterreichischen Partei erwarten Sie sich, daß sie in Zukunft am ehesten etwas tut? Schaffung von mehr Wohnungen, die man sich leisten kann.....“ Die Forderung, daß die ersten Fragen eher neutral gestellt sein sollten, hat SPECTRA also besser erfüllt.

IFES hat eine Reihe von Fragenbatterien verwendet, um möglichst viele Stimuli abfragen zu können. Dies führt, wie schon erwähnt, zur Nachlässigkeit bei Interviewer und Befragten, SPECTRA hat meist Einzelfragen formuliert.

Der sonstige Aufbau des Fragebogens ist bei beiden Instituten ähnlich. Zuerst wurden allgemeine Fragen zum Eindruck von den Parteien gestellt, dann zu den Eindrücken von den Spitzenkandidaten, und relativ am Ende die Sonntagsfrage. SPECTRA hat vor der Sonntagsfrage noch einige Fragen zur Beteiligung an der Wahl gestellt, IFES am Beginn einige Fragen zu bundes- und landespolitischen Themen. Die Formulierungen bei manchen Fragen, die das gleiche Thema behandelten, waren teilweise verschieden, auf dies wird später noch eingegangen.

Es wurden immer geschlossene Fragen gestellt, SPECTRA hat einige Male auch interne Antwortvorgaben verwendet.

„Keine Angaben“-, „Kann nicht sagen“- und „Weiß nicht“- Antwortmöglichkeiten wurden von beiden Instituten verwendet. Antwortkarten wurden nicht verwendet, da die Befragungen ja meist telefonisch durchgeführt wurden.

Tabelle 3.5 zeigt die Länge der Frageneinleitungen, Stimuli und die Anzahl der Antwortkategorien.

Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Tabelle 3.5: Länge der Frageneinleitungen, Stimuli und Anzahl der Antwortkategorien

IFES				SPECTRA			
Frage	Einleitung	Stimulus	Antwortkat.	Frage	Einleitung	Stimulus	Antwortkat.
1	27	7	6	1	17	3	3
2	19	2	4	2	20	3	4
3	18	2	4	3	18	2	3
4	21	3	3	4	22	2	3
5	23	11	2	5	23	3	3
6	30	2	5	6	36	3	4
7	28	4	3	7	18	3	4
8	37	14	5	8	18	3	4
9	12	1	6	9	18	3	4
11	29	11	3	10	18	3	4
12	12	4	6	11	23	3	7
13	11	2	5	12	23	1	7
14	14	1	6	13	16	1	7
15	21	3	5	14	16	1	7
16	26	3	2	15	17	1	7
17	11	3	2	16	13	1	7
18	10	4	7	17	15	1	7
19	8	4	7	18	15	1	7
20	8	4	7	19	4	1	2
21	36	2	7	20	17	3	4
22	10	1	8	21	14	3	7
23	10	1	8	22	15	1	3
24	10	1	8	23	18	2	3
25	10	1	8	24	17	2	3
26	10	1	8	25	44	1	8
				26	14	2	3
				27	24	4	8
				28	42	4	8
Summe	451	92	135		555	61	141
Schnitt	18,04	3,68	5,4		22,2	2,44	5,64

Die durchschnittliche Länge der Frageneinleitungen lag im Schnitt bei SPECTRA etwas über dem Niveau von IFES, aber liegt noch im Bereich von kurzen Frageneinleitungen (bis 24 Wörter), der als optimal anzusehen ist.

Die durchschnittliche Länge des Stimuli lag bei SPECTRA mit 2,44 deutlich unter dem Wert von IFES. Der optimale Bereich liegt bei vier bis dreizehn Worten. Da natürlich oft nur der Name einer Partei ein Stimuli war, ist die Kürze nachvollziehbar.

Trotzdem ist in diesem, für das Meßergebnis sehr bedeutendem Design, die IFES-Vorgangsweise als besser zu beurteilen.

Durchschnittlich wurden etwa fünf bis sechs Antwortkategorien vorgegeben. Der optimale Bereich liegt zwischen vier und sieben. Dieser Bereich wurde von beiden Instituten meist eingehalten, die Anzahl war natürlich durch die Zahl der kandidierenden Parteien oft vorgegeben.

### c) Ergebniss und Hypothesen

Tabelle 3.6: Im Juli durchgeführte Befragungen von IFES und SPECTRA

Partei	IFES/ telefonisch (n=423)	SPECTRA/persönlich (n=547)	SPECTRA/telefonisch (n=1500)
ÖVP	41	41	40
SPÖ	31	26	30
FPÖ	18	23	23
Grüne	8	8	4
LiF	1	2	3
Andere	1	-	-

Das Ergebnis der beiden Institute weicht teilweise deutlich voneinander ab, da aber die Befragungszeiträume verschieden waren, kann man aber keine konkreten Aussagen treffen. Interessant ist allerdings, daß die beiden Umfragen von SPECTRA teils erhebliche Unterschiede aufweisen, dies ist auf die unterschiedlichen Befragungsarten zurückzuführen. Ist dieser Unterschied nun signifikant oder nur rein zufällig. Dies kann man mit dem sogenannten Chi-Quadrat-Test ( $\chi^2$ -Test) überprüfen.

Voraussetzung für diesen Test ist, daß<sup>69</sup>

- Die Werte in absoluten Häufigkeiten vorliegen
- Die Stichproben unabhängig voneinander sind
- Der Erwartungswert in einer Klasse 5 nicht unterschreitet

---

<sup>69</sup> Vgl. [Karmasin/Karmasin 1977] S. 283

Beim  $\chi^2$ -Test wird in den beobachteten Klassen das Quadrat der Differenz von Erwartungswert und beobachteten Wert in Verhältnis zum Erwartungswert gesetzt. Dieser Wert wird durch alle Klassen aufsummiert und mit dem Tabellenwert verglichen, der durch das Signifikanzniveau und der Zahl der Freiheitsgrade  $k$  vorgegebenen ist.

Ist  $T = \sum (h_i - e_i)^2 / e_i > \chi^2_{k-1; 1-\alpha}$ , dann ist die Nullhypothese „Der Unterschied ist rein zufälliger Natur und nicht signifikant“, zu verwerfen und die Unterschiede sind nicht zufälliger Natur.

Die Nullhypothese  $H_0$  lautet in diesem Fall: „Die gefundenen Unterschiede zwischen der telefonischen und der persönlichen Umfrage sind rein zufälliger Natur.“ Zur Ermittlung müssen die Prozentwerte in absolute Werte umgerechnet werden.

Tabelle 3.7a: Beobachtete Häufigkeiten bei den Umfragen von SPECTRA im Juli 97

Partei	Persönlich	Telefonisch	Summe
ÖVP	224	600	824
SPÖ	142	450	592
FPÖ	126	345	471
Grüne	44	60	104
LiF	11	45	56
Summe	547	1500	2047

Den Erwartungswert berechnet man anteilmäßig. Der Anteil der ÖVP-Wähler an den insgesamt 2047 Befragten ist 824. Bei den 1500 Telefoninterviews „erwartet“ man daher  $824/2047 \cdot 1500 = 603,8$  ÖVPler. Somit ergeben sich die Erwartungswerte zu Tabelle 3.7.

Tabelle 3.7b: Erwartete Häufigkeiten

Partei	Persönlich	Telefonisch	Summe
ÖVP	220,1	603,9	824
SPÖ	158,1	433,9	592
FPÖ	126	345	471
Grüne	27,8	76,2	104
LiF	15	41	56
Summe	547	1500	2047

Der Wert der Teststatistik ist dann  $T=28,7$ .

Der Tabellenwert für  $\chi^2_{k-1;1-\alpha}$  ist für vier Freiheitsgrade (fünf Parteien, aber der Wert der fünften Partei ergibt sich aus Gesamtzahl weniger den Werten der anderen vier) und ein Signifikanzniveau von 95%:  $\chi^2_{k-1;1-\alpha} = 9,488$ .

Da 28,7 wesentlich größer ist als 9,488, ist die Nullhypothese zu verwerfen, es besteht also ein signifikanter Unterschied in der Befragungsart.

Interessant ist, daß vor allem die Grünen bei der Telefonumfrage schlechter, und die SPÖ besser abschneiden. Da das Wählerpotential der Grünen in der eher gebildeteren Bevölkerungsschicht, und das der SPÖ in der eher niedrig gebildeten liegt, kann man davon ausgehen, daß sozial Bessergestellte durch Geheimnummern in größerer Zahl nicht erreicht werden als sozial Schwache, die gar kein Telefon besitzen.

Tabelle 3.8: Anfang September durchgeführte Befragungen von IFES und SPECTRA

Partei	IFES/ telefonisch (n=705, 1.9-4.9)	SPECTRA/persönlich (n=547, 1.9-7.9)
ÖVP	42	41
SPÖ	30	28
FPÖ	19	20
Grüne	6	7
LiF	2	4
Andere	1	-

Die Ergebnisse weichen wieder nur minimal voneinander ab. Da diesmal aber die Befragungszeiträume fast übereinstimmen, kann man überprüfen, ob der Unterschied der beiden Institute signifikant ist oder nicht.

H0: „Der Unterschied zwischen der Befragung von IFES und der von SPECTRA ist rein zufälliger Natur.“ Bei der Betrachtung werden nur die fünf Parteien herangezogen, da SPECTRA keinen „Andere“-Prozentsatz auswies.



## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

---

Die Daten von IFES beziehen sich nur auf die 575 Deklarierten minus der „Andere“-Deklarierte.

Tabelle 3.9: Beobachtete Häufigkeiten bei den Umfragen von SPECTRA und IFES Anfang Sept. 97

Partei	IFES	SPECTRA	Summe
ÖVP	242	205	447
SPÖ	173	140	313
FPÖ	109	100	209
Grüne	35	35	70
LiF	12	20	32
Summe	571	500	1071

Tabelle 3.10: Erwartete Häufigkeiten

Partei	IFES	SPECTRA	Summe
ÖVP	238,3	208,7	447
SPÖ	166,9	146,1	313
FPÖ	111,4	97,6	209
Grüne	37,3	32,7	70
LiF	17,1	14,9	32
Summe	571	500	1071

Der Wert der Teststatistik ist  $T=4,21$ . Dies ist geringer als der Tabellenwert von 9,488. Wenn man nur von einem 90%igen Signifikanzniveau ausgeht, so ist der Tabellenwert für  $\chi^2_{k-1;1-\alpha} = 7,779$ . Auch dies ist größer als der berechnete Wert, die Nullhypothese wird angenommen, es gibt keinen signifikanten Unterschied zwischen den Umfragen.

Tabelle 3.11: Letzten vor der Wahl durchgeführte Befragungen von IFES und SPECTRA

Partei	IFES/ telefonisch (n=600, 24.9-28.9)	SPECTRA/persönlich (n=1000, 8.9-18.9)
ÖVP	41	44
SPÖ	29	27
FPÖ	21	19
Grüne	7	7
LiF	1	3
Andere	-	-

Interessant wäre nun, ob gewisse Trends erkennbar sind. Betrachtet werden die Abwärtsbewegung der SPÖ von 30 auf 29% zwischen den beiden letzten IFES-Befragungen, und die Aufwärtsbewegung der ÖVP von 41 auf 44% zwischen den

beiden letzten SPECTRA-Befragungen. Bei den IFES-Daten werden nur die 482 Deklarierten herangezogen.

Die Nullhypothese lautet: „Die Unterschiede der SPÖ-Deklarierten zwischen den beiden Befragungen sind zufälliger Natur“

Tabelle 3.12: Beobachtete SPÖ-Häufigkeiten bei den letzten beiden IFES-Umfragen

	1 Monat vor Wahl	1 Woche vor Wahl	Summe
SPÖ	173	140	313
Nicht-SPÖ	402	342	744
Summe	575	482	1057

Tabelle 3.13: Erwartete SPÖ-Häufigkeiten

	1 Monat vor Wahl	1 Woche vor Wahl	Summe
SPÖ	170,3	142,7	313
Nicht-SPÖ	404,7	339,3	744
Summe	575	482	1057

Der Wert der Teststatistik ist  $T=0,133$ . Verglichen mit dem Tabellenwert (für einen Freiheitsgrad) von 3,841, kann man die Nullhypothese annehmen. Ein signifikanter Trend ist nicht beobachtbar.

Die Nullhypothese für die ÖVP-Daten ist gleichlautend: „Die Unterschiede der ÖVP-Deklarierten zwischen den beiden Befragungen sind zufälliger Natur“

Tabelle 3.14: Beobachtete ÖVP-Häufigkeiten bei den letzten beiden SPECTRA-Umfragen

	1.9. bis 7.9.	8.9. bis 18.9.	Summe
ÖVP	205	440	645
Nicht-ÖVP	295	560	855
Summe	500	1000	1500

Tabelle 3.15: Erwartete ÖVP-Häufigkeiten

	1.9. bis 7.9.	8.9. bis 18.9.	Summe
ÖVP	215	430	645
Nicht-ÖVP	285	570	855
Summe	500	1000	1500

Der Wert der Teststatistik ist  $T=1,22$ . Obwohl die Änderung von 41 auf 44% relativ groß erscheint, ist sie also bei weitem nicht signifikant.

Das Erkennen von Trends kann also nicht auf statistisch abgesicherten Werten beruhen, sondern erfordert Fingerspitzengefühl und Erfahrung der Institute.

Abschließend soll noch untersucht werden, welches Institut das tatsächliche Ergebnis genauer vorhergesagt hat.

Tabelle 3.16: Die letzten Ergebnisse verglichen mit dem tatsächlichen Wahlergebnis

Partei	Ergebnis	IFES	d	SPECTRA	d
ÖVP	42,69	41	1,69	44	1,31
SPÖ	27,04	29	1,96	27	0,04
FPÖ	20,63	21	0,37	19	1,63
Grüne	5,78	7	1,22	7	1,22
LIF	2,09	1	1,09	3	0,91
Andere	1,77	-	-	-	-
TE			6,33		5,11
MAE			1,27		1,02
RSME			1,38		1,16

Das Ergebnis von SPECTRA ist also marginal besser wie jenes von IFES. Auffällig ist das beide Institute, ihren jeweiligen Auftraggeber etwas deutlicher über den tatsächlichen Wahlergebnis prognostiziert haben.

#### d) Weitere Fragestellungen

Bei den Umfragen von IFES und SPECTRA sind eine Reihe von wichtigen Fragen gestellt worden, die für strategische Überlegungen wichtig sind. Hier werden drei wesentliche behandelt, und zwar der Eindruck der Befragten von den Spitzenkandidaten der Parteien, die „Lambach“-Frage und die „Straßenbauaffäre“-Frage, die beiden letzteren wurden von IFES gestellt.

Bei der Frage nach den Spitzenkandidaten wurde von beiden Instituten zunächst die „Kennen Sie...?“-Frage gestellt. Die Beliebtheit der Politiker wurde bei IFES mit „Und haben Sie von ..... ganz allgemein eine günstige Meinung?“ abgefragt, bei SPECTRA mit „Welchen Eindruck haben sie derzeit von .....? Eher einen guten oder eher keinen guten Eindruck?“.

## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

---

Tabelle 3.17: Positive Meinung von den Spitzenkandidaten bei IFES und SPECTRA

Politiker	IFES Juli	SPECTRA Juli	IFES 1.9.-4.9.	SPECTRA 12.9-17.9.	IFES 24.9.-28.9.
Josef Pühringer	67	68	67	74	76
Fritz Hochmair	46	50	49	52	40
Hans Achatz	33	32	31	32	28
Rudi Anschober	33	36	28	48	32
Christa Fischer-Korp	5	25	8	25	8

Während bei den Spitzenkandidaten der drei großen Parteien die Daten relativ übereinstimmen, gehen sie bei Rudi Anschober und Christa Fischer-Korp doch weit auseinander. So haben 25% eher einen guten Eindruck von der liberalen Spitzenkandidatin, aber nur 5-8% eine günstige Meinung. Nicht ganz so stark ist der Unterschied beim grünen Spitzenkandidaten.

IFES fragte auch noch, welche Person der Befragte direkt zum Landeshauptmann wählen würde, hier war nur eine Nennung möglich, diese Frage wurde allerdings Ende September nicht mehr gestellt. SPECTRA fragte: „Von welchem Politiker haben sie den relativ besten Eindruck?“

Tabelle 3.18: Besten Eindruck bzw. Landeshauptmannpräferenz

Politiker	IFES Juli	SPECTRA Juli	IFES 1.9.-4.9.	SPECTRA 12.9-17.9.
Josef Pühringer	51	52	55	57
Fritz Hochmair	17	13	16	13
Hans Achatz	9	7	10	5
Rudi Anschober	5	7	3	8
Christa Fischer-Korp	-	1	2	1
Keinen	6	8	5	6
Weiß nicht	12	12	10	10

Hier stimmen die Daten noch genauer überein. Die Unterschiede sind zwar nicht groß, aber es fällt trotzdem auf, daß

- sich etwas mehr für Fritz Hochmair als Landeshauptmann entscheiden würden, obwohl sie nicht den besten Eindruck von ihm haben. Dies könnte man als Loyalität einiger SPÖ-Wähler deuten.

## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

- sich etwas mehr für Hans Achatz als Landeshauptmann entscheiden würden, obwohl sie nicht den besten Eindruck von ihm haben. Dies könnte man als Loyalität einiger FPÖ-Wähler deuten.
- viele von Rudi Anschober die beste Meinung haben, ihn aber trotzdem nicht zum Landeshauptmann wählen würden.

Die Frage der Errichtung des Kraftwerkes Lambach war ein sehr heikles Thema im Winter 96/97 in Oberösterreich. Die ÖVP war vehement dafür, die SPÖ hat sich nach jahrelangem Hin und Her auf die Seite der Kraftwerksgegner geschlagen. Nun wollte man von SPÖ-Seite her natürlich wissen, ob man dieses Thema im Wahlkampf gegen den Hautgegner ÖVP verwenden sollte.

Abbildung 3.7: IFES-Umfrage im Juli 97: Was die Errichtung des Kraftwerkes bei Lambach betrifft, sind Sie da....?

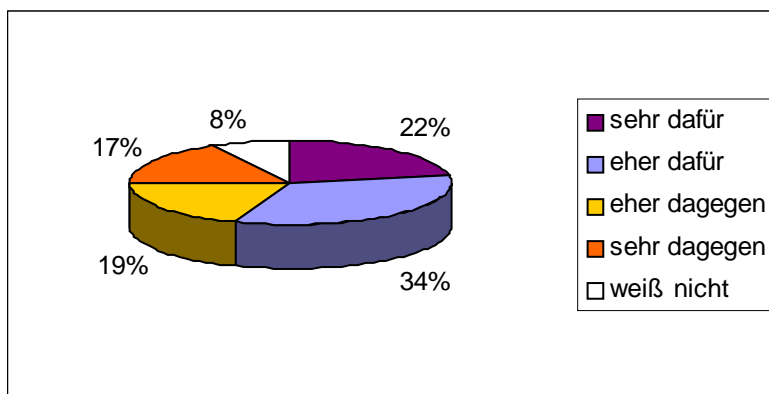
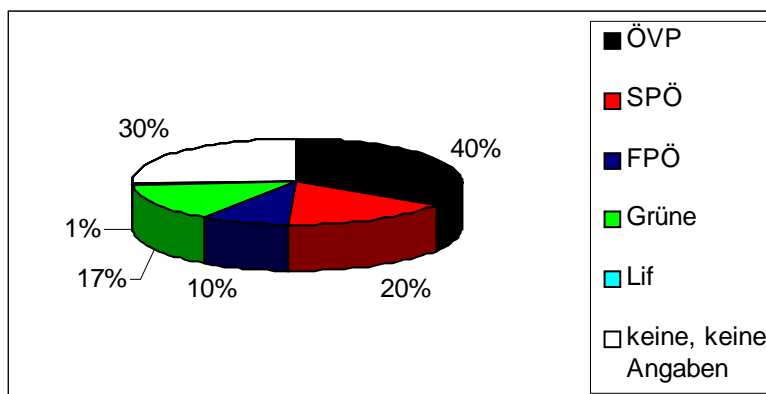


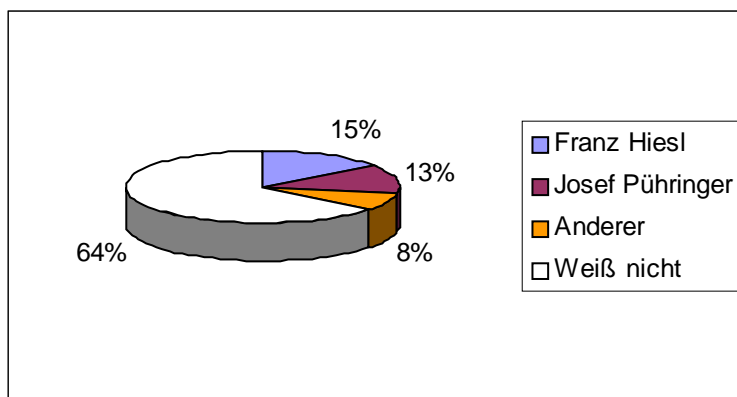
Abbildung 3.8: Welche Partei hat im Konflikt um Lambach die gleiche Position wie sie selbst?



Die Bevölkerung war also mehrheitlich für die Errichtung des Kraftwerk Lambachs. Auch die Befragten, die angaben, SPÖ wählen zu wollen, waren mehrheitlich für Lambach. Mehrheitlich stand die Bevölkerung auf der Seite der ÖVP, auch die deklarierten SPÖ-Wähler waren nur schwach auf der SPÖ-Seite. Profitiert vom Lambach-Konflikt haben die ÖVP und die Grünen. So war klar, daß dies kein Thema für den Wahlkampf der SPÖ war.

Ein weiteres Konfliktthema, aber Jahre zuvor, war die sogenannte Straßenbauaffäre, bei der es zu Preisabsprachen gekommen ist. Dadurch wurde zu überhöhten Preisen gebaut. Josef Pühringer war damals Bau-Landesrat, er übergab diesen Job an Franz Hiesl, als er Landeshauptmann wurde. Bei der Umfrage von IFES gaben 298 von den 423 Befragten an, von der Bauaffäre bereits gehört oder gelesen zu haben, das sind immerhin 71%. Diese 71% wurden befragt, welcher Politiker für die Affäre verantwortlich ist.

Abbildung 3.9: Und welcher Politiker ist ihrer Meinung nach für diese Affäre hauptverantwortlich?



Die Befragten wußten zwar von der Affäre, hatten aber keine Ahnung, in wessen Kompetenzen diese Sache fiel. Damit war auch dieser Angriffspunkt für die SPÖ vom Tisch, die dann eine sehr moderate und erfolglose Wahlkampflinie einschlug.

## IV. Vergleich der beiden Methoden

### IV.1 Theoretischer Hintergrund

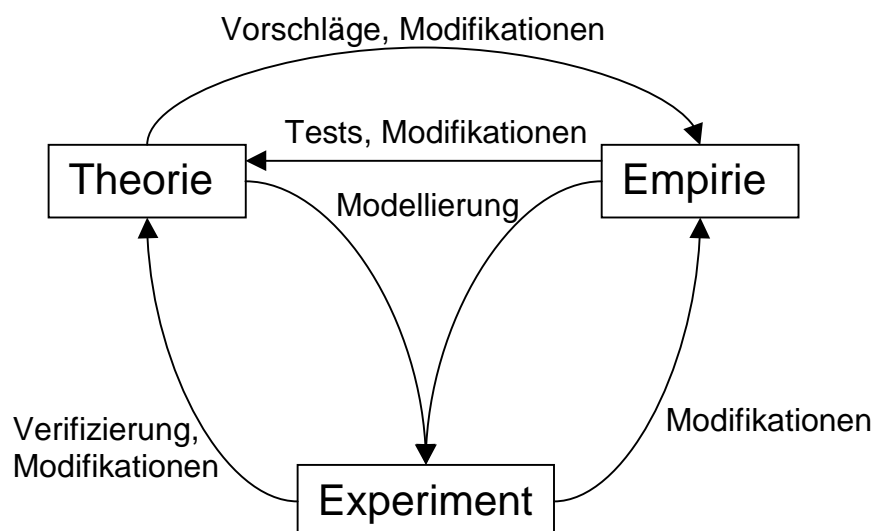
Im Abschnitt III.1.1 wurden die Methoden der empirischen Sozialforschung vorgestellt, in der die Umfrageforschung eingebettet ist.

Ein Political Stock Market ist ein Experiment, der theoretische Hintergrund ist die experimentelle Ökonomie.

Experimente wurden in der Wissenschaft bis vor ca. 400 Jahren nicht angewandt. Ausgehend von Physik und Chemie wurden viele andere wissenschaftliche Disziplinen durch experimentelle Techniken revolutioniert.

Die Ökonomie wurde in den letzten 40 Jahren eine experimentelle Wissenschaft, der Kreislauf zwischen Theorie und Empirie wurde durch die Experimentelle Forschung erweitert.

Abbildung 4.1: Experimente in den Wirtschaftswissenschaften <sup>70</sup>



---

<sup>70</sup> Vgl. Ortner [1996] S. 8

Wie in den Naturwissenschaften werden Ideen zuerst in Labor- oder Feldexperimenten entwickelt, erprobt und verbessert, bevor sie in realen Projekten zur Anwendung kommen. Prognosesysteme auf Marktbasis wie Political Stock Markets sind ein gutes Beispiel dafür, die zuerst in einfachen Laborversuchen über global laufende Feldexperimente bald auch Einzug in industrieller oder volkswirtschaftlicher Problemstellungen (Prognose von Kennzahlen) finden können. Auf industrielle Anwendungen wird in Kapitel IV.4 noch näher eingegangen.

Eine Kategorisierung von Experimenten kann nach der Art des Experiments in Labor- oder Feldexperiment, und nach dem Umweltzustand in kontrollierte bzw. unkontrollierte Umweltbedingungen getroffen werden. Dadurch ergeben sich vier mögliche Kombinationen. Bei einem PSM handelt es sich um ein kontrolliertes Feldexperiment, das Regelwerk wird vorgegeben und in die reale Welt ausgesetzt. Der Experimentator kann aber regelnd eingreifen und im Extremfall das Experiment auch abbrechen.

## **IV.2 Kostenvergleich**

Die Kosten des APSM sind schwer einzuschätzen, da die benötigte Hardware und das Netzwerk natürlich auch anderweitig verwendet werden und die Software für mehrere Projekte eingesetzt werden kann. Die hier aufgestellten Kosten sind daher nur als grobe Abschätzung zu betrachten.

- 1) Hardware: Benötigt wird ein PC mit ausreichender Datenverarbeitungsgeschwindigkeit, die einmalige Anschaffung kostet 20.000 öS. Da dieser PC natürlich anderweitig verwendet werden kann, wird für das OÖ-Projekt 20% angesetzt, das heißt, 4.000 öS.
- 2) Netzwerk: Benötigt wird ein permanenter Internetzugang (Standleitung), die Kosten betragen ca. 10.000 öS pro Monat. Das OÖ-Projekt lief ca. drei Monate, auch hier werden 20% angesetzt, die Kosten ergeben sich damit zu 6.000 öS.



- 3) Software: Die Kosten werden auf 50.000 öS geschätzt, da die Software für mehrere Projekte verwendet werden kann, werden 25%, also 12.500 öS angesetzt.
- 4) Betreuung des Marktes: Geschätzt wird eine Arbeitsstunde pro Tag, die Kosten einer Arbeitsstunde werden mit 700 öS angesetzt, die Dauer des Projekts auf 90 Tage, damit ergeben sich 63.000 öS.

Die Kosten akkumulieren sich zu 85.500 öS, wobei die Betreuung des Marktes über 70% der Kosten ausmacht, dadurch wirken sich Fehleinschätzungen in diesem Bereich gravierend auf die Gesamtkosten aus.

Bei Meinungsumfragen sind die Anzahl der Interviews der gravierendste Kostenfaktor. Der Detaillierungsgrad des Fragebogens und der Aufwand zur Erstellung dieser wirken sich nicht beträchtlich aus. Auch die Interviewlänge hat keinen großen Einfluß auf die Kosten, da die zeitunabhängigen Kosten für ein Interview (Aussuchen der Zielperson, kontaktieren, besuchen,...) den Hauptanteil des Aufwandes einnimmt.

Die genauen Kosten der Projekte für die Oberösterreichische Landtagswahl wurden weder von IFES noch von SPECTRA angegeben. Laut Information von IFES kann als Faustformel 180 bis 260 öS pro Interview angesetzt werden. Dies sind die Gesamtkosten einer Untersuchung bezogen auf die Interviewanzahl. Die Kosten bei den Untersuchungen zur Oberösterreichischen Landtagswahl werden mit 200 öS pro Interview angenommen.

Im gleichen Zeitraum wie das APSM-Projekt lief wurden von IFES drei Befragungen durchgeführt mit 423, 705 bzw. 600 Interviews. Dadurch ergeben sich die Kosten zu  $(423+705+600)*200 = 345.600$  öS.

Die Kosten dieser Meinungsumfragen sind also um einen Faktor 4 höher als beim APSM-Projekt.

Da SPECTRA wesentlich mehr Interviews durchgeführt hat, werden die Kosten dieses Instituts beträchtlich höher ausgefallen sein. Da die Genauigkeit des

Ergebnisses aber letzten Endes nicht wesentlich besser war, kann man daraus schließen, daß die Strategie von IFES (weniger Interviews, aber genaue Planung des Fragebogens zur Reduzierung der Fehler, die in III.3 besprochen wurden) eine kostengünstige, unter Beibehaltung ausreichender Genauigkeit, darstellt.

### IV.3 Ergebnisse

Die Tabellen 2.7 und 3.16 werden hier noch einmal zusammengefaßt.

Tabelle 4.1: Vergleich der Ergebnisse von APSM und letzten Meinungsumfragen

Partei	Ergebnis	APSM	d	IFES	d	SPECTRA	d
ÖVP	42,69	40	2,69	41	1,69	44	1,31
SPÖ	27,04	31	3,96	29	1,96	27	0,04
FPÖ	20,63	22	1,37	21	0,37	19	1,63
Grüne	5,78	7	1,22	7	1,22	7	1,22
LIF	2,09	4	1,91	1	1,09	3	0,91
Andere	1,77	-	-	-	-	-	-
TE			11,15		6,33		5,11
MAE			2,23		1,27		1,02
RSME			2,45		1,38		1,16

Die Meinungsumfragen haben bei der Oberösterreichischen Landtagswahl klar die genaueren Ergebnisse prognostiziert. Der Hauptgrund ist, wie schon erwähnt, die geringe Teilnahme am APSM97-Projekt und die Dominanz von wenigen Teilnehmern. Ein weiterer Grund besteht darin, daß es für Meinungsforschungsinstitute wesentlich leichter ist, Wahlprognosen zu Landtagswahlen zu erstellen, da die Veränderungen in den letzten Wochen nicht so gravierend sind wie auf Bundesebene, das heißt, das Wählerverhalten wesentlich stabiler ist.

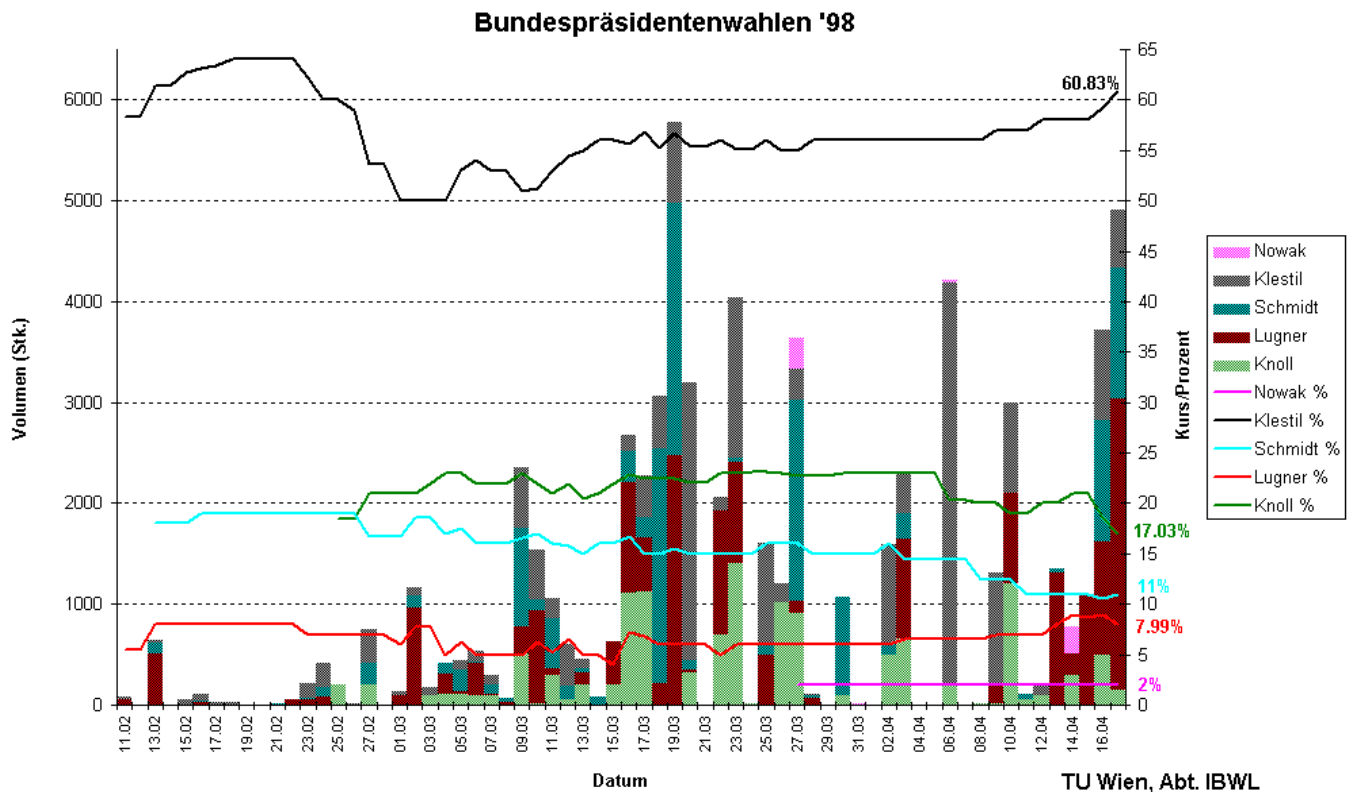
Bei der Bundespräsidentenwahl im April 1998 in Österreich zeigte sich aber wieder, welch beeindruckendes Ergebnis ein PSM erzielen kann, obwohl auch bei diesem Projekt das Interesse mit insgesamt nur 25 aktiven Tradern relativ gering war.

## Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Tabelle 4.2: APSM versus Meinungsforschung – Bundespräsidentenwahl 1998

Kandidat	Ergebnis	APSM	d	Gallup	d
Klestil	63,49	60,83	2,66	66	2,51
Schmidt	11,08	11	0,08	10	1,08
Knoll	13,53	17,03	3,5	16	2,47
Lugner	9,94	7,99	1,95	6	3,94
Nowak	1,95	2	0,05	1	0,95
TE			8,24		10,95

Abbildung 4.2: Kursverläufe und Handelsvolumen beim BP-Wahlmarkt



Von 1994 bis 1998 wurden zahlreiche APSM-Projekte durchgeführt, mit unterschiedlichem Erfolg. Tabelle 4.3 vergleicht die Ergebnisse aller durchgeführten APSM-Projekte mit der **jeweils besten** Meinungsumfrage vor der Wahl, wobei als Qualitätsparameter der „Mean Average Error“ und der „Root Square Mean Error“ herangezogen wird.

Tabelle 4.3: MAE und RSME bei den verschiedenen Projekten

Wahl	APSM		Umfrage	
	MAE	RSME	MAE	RSME
NRW94	1,18	1,35	1,48	1,74
NRW95	2,71	3,76	2,19	3,17
Stmk.	2,97	4,27	2,35	2,89
Wien	1,15	1,47	1,27	1,59
EU	2,29	2,56	1,98	2,3
OÖ	2,23	2,45	1,02	1,16
Bundespr.	1,65	2,15	2,19	2,45
Mittelwert	2,05	2,69	1,78	2,19
Standardab.	0,719	1,102	0,521	0,725

Im Durchschnitt lagen der MAE und RSME bei den **besten** Umfragen unter den Werten der APSM-Projekte. Der Mittelwert des MAE liegt bei den APSM-Projekten nur 15% über dem der Umfragen, vernachlässigt man die beiden Wahlen mit äußerst schwacher Beteiligung – Oberösterreich und Steiermark – so sind es nur 1%, die Ergebnisse können also als gleich angesehen werden.

Die Situation beim RSME ist allerdings eine andere, der Mittelwert bei den APSM-Projekten liegt 23% über dem der Umfragen. Selbst bei Vernachlässigung der OÖ- und Steiermark-Wahl verbleiben über 7%. Das heißt, daß bei den APSM-Projekten die Werte einzelner Parteien stärker abweichen als bei den Umfragen.

Im Hinblick auf die Standardabweichung zeigt sich auch, daß die Meinungsumfragen eindeutig die stabileren Ergebnisse liefern, daß heißt, die Schwankungen um den Mittelwert sind beim APSM größer. Die Unsicherheit der Fehlergröße ist bei den Umfragen geringer, dies geht auch aus den Abbildungen 4.1 und 4.2 hervor.

# Vergleich von Wahlprognosen durch Meinungsumfragen und Bewertung durch Börsen anhand der Oberösterreichischen Landtagswahl 1997

Abbildung 4.3: Mean Average Error bei APSM und Umfragen

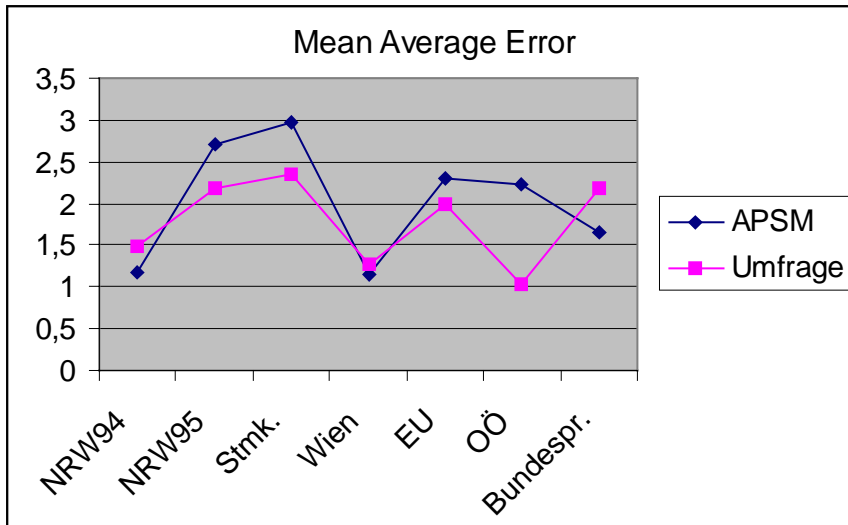
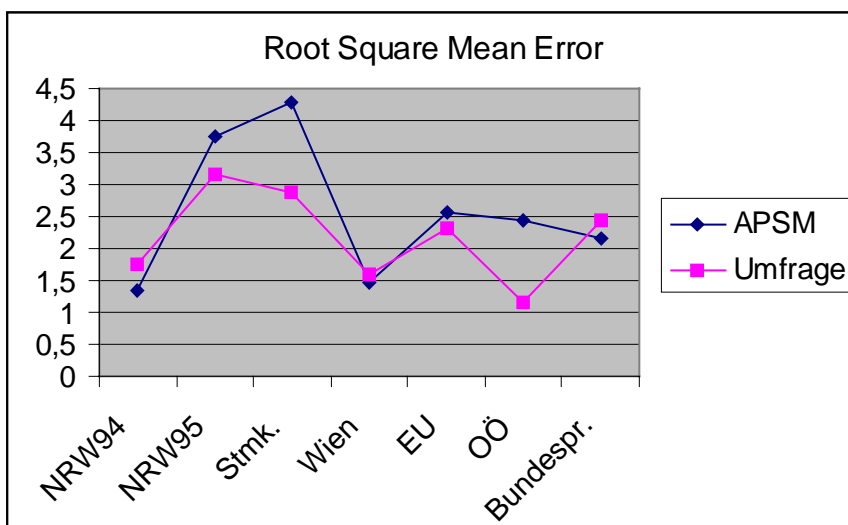


Abbildung 4.4: Root Square Mean Error bei APSM und Umfragen



Der APSM konnte also bei der NRW94, der Wien-Wahl und bei der Bundespräsidentenwahl die genaueren Ergebnisse erzielen. Sieht man von der OÖ- und der Steiermark-Wahl ab, so verbleiben die NRW95 (Markt wurde allerdings manipuliert) und die EU-Wahl, bei der die besten Meinungsumfragen die besseren Resultate erzielen konnten.

## **IV.4 Zukünftige Perspektiven**

### ***IV.4.1 Prognosen für politische Wahlen***

Bei den zahlreich durchgeführten PSM-Projekten konnten teilweise beachtliche Ergebnisse erzielt werden. PSM weisen einige Vorteile gegenüber Meinungsumfragen auf:

- Repräsentativität der Teilnehmer nicht gefordert
- Kontinuierliche Bewertung der Aktientitel und damit kontinuierliche Prognose im Vergleich zu Umfragen (Trenduntersuchungen notwendig)
- Anonyme Teilnahme, einfaches Design, geringer Aufwand und deshalb
- Geringere Kosten.

Der Hauptnachteil liegt in der Unsicherheit der Ergebnisqualität, da der Erfolg von Parametern abhängt, die größtenteils vom Experimentator nicht beeinflusst werden können.

Für politische Parteien sind Meinungsumfragen unersetzlich, da sie vor allem für strategische Überlegungen wertvolle Informationen liefern können. Die Genauigkeit der Prognose ist für politische Parteien nicht so entscheidend, da vor allem das Erkennen von Trends und Strömungen von Bedeutung ist.

Das möglichst genaue Vorhersagen von Wahlen ist hauptsächlich mediales Interesse. Somit wird die Zukunft des PSM und des APSM auch weiterhin darin liegen, in Zusammenarbeit mit Medien, Kursentwicklungen und Prognosen der Öffentlichkeit zur Verfügung zu stellen.

#### **IV.4.2 Prognosen von Kennzahlen und in Industrie**

Die Vorteile der Prognose durch Bewertung durch Börsen können natürlich auch für andere Problemstellungen genutzt werden. Eine interessante Möglichkeit ist das Prognostizieren von volkswirtschaftlichen Kennzahlen wie Arbeitslosenrate, Inflationsrate,...

Hier soll allerdings etwas näher auf die Möglichkeit der innerbetrieblichen Anwendung eingegangen werden, konkret auf ein Softwareprojekt von Siemens im Telekommunikationsbereich. Bei diesem Projekt wurde ein firmeninterner Aktienmarkt als Projektmanagement Tool verwendet.

Eines der Hauptprobleme im Projektmanagement ist das Projekt-Controlling, die rechtzeitige Fertigstellung von Projekten ist äußerst wichtig, Verspätungen sind meist mit großen internen und externen Kosten verbunden. Es gibt einige Projektmanagementmethoden, die versuchen, Verspätungen so schnell wie möglich zu erkennen um ein rechtzeitiges Gegensteuern zu ermöglichen.

Aber besonders Schwierigkeiten wie technische Probleme, Verzögerungen,... werden oft nur sehr langsam oder gar nicht an das Projektmanagement weitergeleitet, weil die involvierten Personen befürchten, Konsequenzen des negativen Feedbacks tragen zu müssen.<sup>71</sup>

Das heißt, ein Vorhersagemarkt, der die Informationen und Einschätzungen der involvierten Personen aggregiert, könnte dem Management zusätzliche, anonyme Informationen liefern. Jeder Projektmitarbeiter sammelt Informationen aus seinem eigenen Umfeld über das Vorankommen des Projektes und bringt somit kleine Informationsbruchstücke in den Markt ein.

Durch einfache Beobachtung der Preisentwicklung kann abgeschätzt werden, ob mit einer rechtzeitigen Fertigstellung des Projektes zu rechnen ist, oder wie lange sich das Projekt verzögern wird.

---

<sup>71</sup> vgl. [Ortner 1997/1998] S. 2

Eine gezielte Umfrage unter den Projektteilnehmern wäre in diesem Falle auch nicht hilfreich, da sich Veränderungen äußerst rasch ergeben können und daher die Reaktionsgeschwindigkeit dieser Methode nicht ausreichend ist. Außerdem wäre die Gefahr von verzerrenden Aussagen der Projektmitarbeiter aus Angst vor Konsequenzen zu groß.

### Siemens-Projekt

In einem Softwareprojekt im Telekommunikationsbereich waren etwa 200 Personen, der Großteil war Ingenieure, beschäftigt. Das Projekt sollte innerhalb von sechs Monaten fertiggestellt sein, das Projektende war genau definiert, nämlich wenn ein spezielles Testprogramm an eine andere Abteilung in Deutschland übergeben wurde, das bedeutete Erreichung des End-Meilensteins.

Ein voll computerisierter DAM wurde am Intranetzwerk von Siemens Österreich installiert, die verwendete Software war die von der Kumo Inc. entwickelte.

Der Markt, seine Ziele und der Ablauf des Projektes wurden in einem ersten kurzen Briefing den Projektmitarbeitern vorgestellt. Freiwillige, die bereit waren, 100 öS zu investieren, wurden von Siemens mit weiteren 200 öS motiviert. Über 50 Projektmitarbeiter entschieden sich, am Markt teilzunehmen.

Handeln an zwei verschiedene Märkte war möglich. Der erste Markt fragte simple, ob die Fertigstellung des Projektes zeitgerecht erfolgen kann. Die Auszahlungsregel war Winner Takes All. Für einen „Ja“-Anteil wären bei rechtzeitiger Fertigstellung am Ende 1 öS ausbezahlt worden. Für einen „Nein“-Anteil wäre 1 öS bei Verzögerung und 0 öS bei rechtzeitiger Fertigstellung ausbezahlt worden.

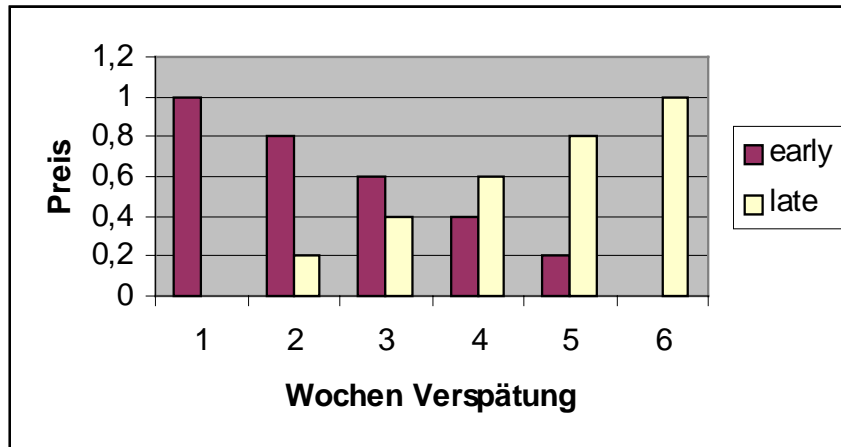
Der zweite Markt („Verzug“) sollt eine mögliche Verzögerung vorhersagen. Ein „Ja“-Anteil oder „Early“ und ein „Nein“-Anteil oder „Late“ hatten folgende lineare Auszahlungsregel:



JA:  $\max(1 - 0,2 \cdot \text{Wochen Verspätung}; 0)$  öS

NEIN:  $\min(0,2 \cdot \text{Wochen Verspätung}; 1)$  öS

Abbildung 4.5: Auszahlungsregel im Verzug-Markt

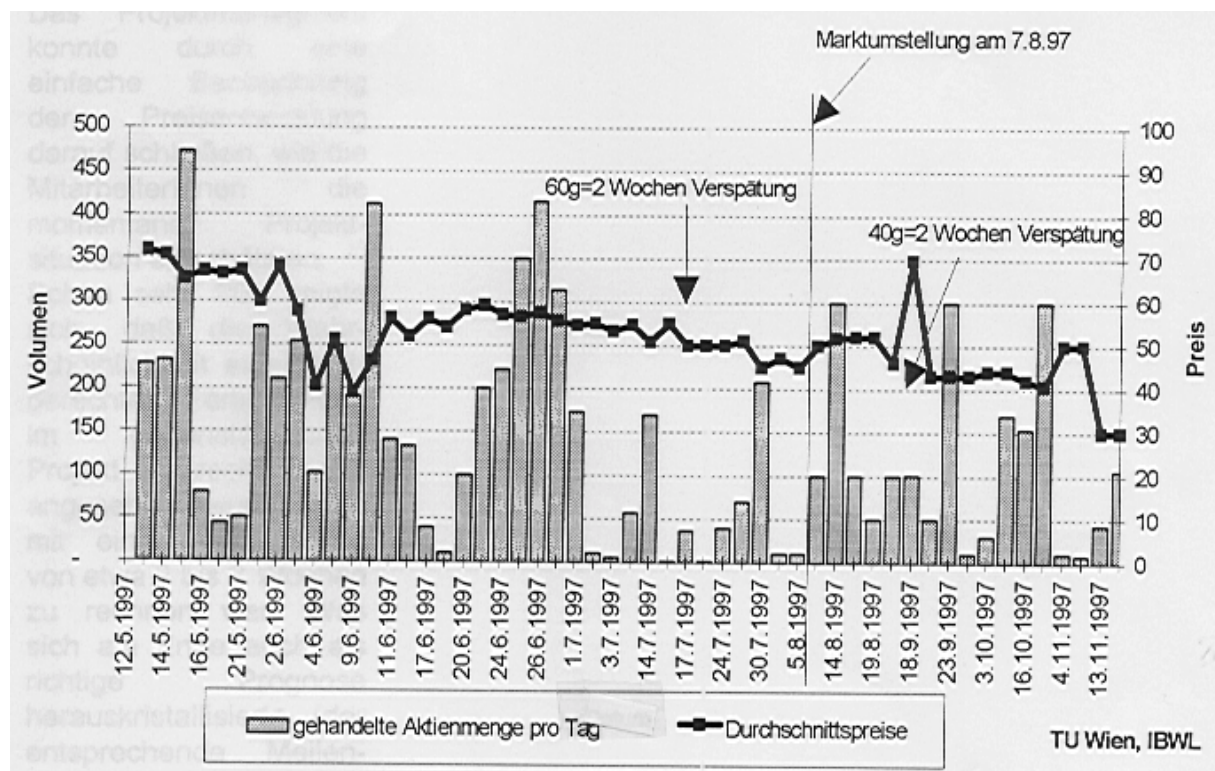


Der Handel begann Mitte Mai 97. Schon frühzeitig zeigte sich, daß mit einer rechtzeitigen Fertigstellung des Projektes nicht gerechnet wurde, die durchschnittlich gehandelten Preise sanken im Winner Takes All-Markt innerhalb von drei Wochen von anfangs 0,9 öS auf 0,4 öS.

Im Verzug-Markt schwankten die Preise anfangs zwischen 0,7 öS und 0,4 öS, diese bedeutet 1,5 bis drei Wochen Verspätung. Ab Mitte Juni pendelten sich der Preis zwischen 0,5 öS und 0,6 öS ein. Am 7.8 wurde der Markt umgestellt, nun entsprach ein Preis von 0,4 öS einer Verzögerung von zwei Wochen im Gegensatz zu 0,6 öS vorher. Die Umstellung war notwendig, da die Terminvorgaben vom Kunden geändert wurden. Mitte September pendelte sich der Preis wieder um 0,4 öS ein, die entspricht einer Verzögerung von zwei Wochen.

Am Ende entsprach dies auch relativ exakt der tatsächlichen Verzögerung, der End-Meilenstein wurde 13 Tage später als geplant erreicht.

Abbildung 4.6: Preisentwicklung und Handelsvolumen im Markt „Verzug“<sup>72</sup>



Zusammenfassend kann gesagt werden, daß das erste Projekt dieser Art durchaus ein Erfolg war. Die Teilnahme an den Marktaktivitäten war von Beginn an überraschend groß, Marktgleichgewicht wurde bereits nach ca. einem Monat erreicht.

Daraus läßt sich schließen, daß das Design des Marktes und die Implementierung des Marktes (Software, Benutzerschnittstelle,...) den Teilnehmern schnell vertraut wurde. Die gewünschten Informationen wurden nach äußerst kurzer Zeit aggregiert.

Man kann also mit dem ersten Einsatz dieses Verfahrens zufrieden sein. Verbesserungen und weitere betriebliche Anwendungen lassen sich diskutieren.

<sup>72</sup> vgl. [Ortner 1997] S. 1

## V. Conclusio

In dieser Arbeit wurden die zwei Möglichkeiten verglichen, Prognosen zu erstellen, und zwar durch Meinungsumfragen und durch Bewertung durch Börsen.

Zuerst wurde das Konzept der Double Auction Märkte vorgestellt, die auch bei Political Stock Markets verwendet werden. Nach einer Vorstellung der kurzen Geschichte der PSM und von wichtigen Begriffen wurde auf das Ergebnis des APSM97 eingegangen.

In diesem Markt wurde eine neue Software verwendet, die sich vor allem auch beim APSM zur Bundespräsidentenwahl 1998 bewährt hat. Die gravierendste Veränderung zur alten Software ist, daß eine Strategie des „offenen Orderbooks“ verfolgt wird. Zukünftige Beobachtungen werden zeigen, ob diese Strategie beibehalten werden soll.

Die Teilnahme am APSM97 war äußerst gering, vor allem die Dominanz zweier Marktteilnehmer wirkte sich negativ auf das Ergebnis aus. Eine große Teilnehmerzahl hat gravierenden Einfluß auf die Qualität des Ergebnisses und ist ein wichtiger Parameter für einen erfolgreichen PSM. Da die Zahl der Internetuser in den nächsten Jahren steigen wird, kann man in Zukunft mit einem größeren potentiellen Teilnehmerkreis rechnen, dadurch wird sich bei gegebenem medialen Interesse an einer Wahl die Teilnehmerzahl bei zukünftigen Projekten vergrößern.

Im zweiten Hauptteil der Arbeit wurde die Umfrageforschung als Methode der empirischen Sozialforschung vorgestellt. Die Stichprobenauswahl spielt bei Umfragen eine wesentliche Rolle, die Strukturierung des Objektbereichs und die verschiedenen Auswahlverfahren wurden vorgestellt. In Österreich verwenden die führenden Meinungsforschungsinstitute ein mehrstufiges Clusterauswahlverfahren.

Das Interview und der Fragebogaufbau wurden erläutert, wie Fragen formuliert werden können und welchen Einfluß dies auf definierte Meßqualitätskriterien wie Validität und Reliabilität hat.

Ein wichtiger Punkt ist, daß der Auftraggeber, sofern es sich um eine politische Partei handelt, hauptsächlich an Informationen interessiert ist, die strategische wichtig sind, eine möglichst genaue Prognose des Wahlergebnisses ist nicht so bedeutend. Der Hauptteil der Fragen beschäftigt sich daher mit Personen, politischen Ereignissen und Einstellungen des Befragten. Die „Sonntagsfrage“ hat nicht die Bedeutung wie bei den PSM, wo es das einzige Ziel ist, das Wahlergebnis möglichst genau zu prognostizieren.

Die Fragebögen der Institute IFES und SPECTRA, die bei der Oberösterreichischen Landtagswahl tätig waren, wurden analysiert, wobei IFES bei der Strukturierung und der Fragenformulierung etwas besser zu bewerten ist. Das Ergebnis von IFES ist trotz wesentlich kleinerer Stichprobeneinheiten nur unwesentlich schlechter als jenes von SPECTRA.

Zahlreiche Hypothesen wurden aufgestellt und mit dem  $\chi^2$ -Test getestet. Hypothesen spielen bei Meinungsumfragen eine große Rolle, da sie gesetzesähnliche Aussagen darstellen, die mit einem gewissen Vagheitsspielraum als bestätigt angesehen werden können. Die getesteten Hypothesen bei den Umfragen von IFES und SPECTRA zeigten, daß sich Veränderungen bei Trenduntersuchungen fast immer innerhalb der statistischen Schwankungsbreite bewegen. Es erfordert daher viel Erfahrung und Fingerspitzengefühl der Institute, Trends erkennen zu können, da die statistische Signifikanz fehlt.

Ein Vergleich der beiden Methoden zeigte, daß die APSM die wesentlich günstigere Variante darstellen. Die Ergebnisse der bisherigen Projekte sind nur marginal schlechter wie die der Meinungsumfragen, obwohl die Unsicherheit der Ergebnisqualität bei den APSM größter ist, vor allem auch, weil der Erfolg von nicht kontrollierbaren Parametern abhängt.

Für politische Wahlen sind Meinungsumfragen unersetzlich, die Rolle der APSM wird weiterhin jene sein, mediales Interesse an genauen Wahlprognosen zu befriedigen.

Große Zukunft für Prognose durch Bewertung durch Börsen zeigt sich aber auch in anderen Gebieten, wie der Prognose von volkswirtschaftlichen Kennzahlen und innerbetrieblichen Vorhersagen.

In dieser Arbeit wurde kurz der Einsatz von Vorhersagemärkten als Projektmanagement-Tool vorgestellt. Die vorhandenen Informationen werden schnell aggregiert, Gleichgewichtspreise stellen sich rasch ein. Das einfache und für die Teilnehmer verständliche Design ermöglicht es, auch mit Aktienmärkten unerfahrene Mitarbeiter in das Experiment zu integrieren.

Die äußerst zufriedenstellenden Ergebnisse im Vorhersagemarkt für ein Softwareprojekt von Siemens im Telekommunikationsbereich, daß das erste Experiment dieser Art darstellte, lassen zuversichtlich auf zukünftige innerbetriebliche Anwendungen blicken.

## VI. Tabellen und Abbildungsverzeichnis

Abbildung 2.1: Induzierte Angebots- und Nachfragekurven _____	6
Tabelle 2.1: Verlauf der Abschlüsse in der fünften Periode _____	7
Abbildung 2.2: Ergebnisse des Experiments _____	7
Tabelle 2.2: APSM versus Meinungsforschung – Nationalratswahl 1994 _____	12
Abbildung 2.3: Kursverlauf des Nationalratwahlmarkt 94 _____	12
Tabelle 2.3: APSM versus Meinungsforschung – Nationalratswahl 1995 _____	13
Abbildung 2.4: Kursverläufe des Nationalratwahlmarkts 95 _____	14
Abbildung 2.5: Handelsvolumen im Nationalratwahlmarkt 95 _____	14
Tabelle 2.4: APSM versus Meinungsforschung – Steirischer LTW 95 _____	15
Abbildung 2.6: Kursverläufe und Handelsvolumen bei der Steirischen LTW 95 _____	15
Tabelle 2.5: APSM versus Meinungsforschung – Wiener GRW 96 _____	16
Abbildung 2.7: Kursverläufe und Handelsvolumen bei der Wiener GRW 95 _____	16
Tabelle 2.6: APSM versus Meinungsforschung – EU-Wahlmarkt 96 _____	17
Abbildung 2.8: Kursverläufe und Handelsvolumen beim EU-Wahlmarkt _____	17
Abbildung 2.9-2.17: Umfrage unter den Teilnehmern _____	31
Abbildung 2.18: Gehandelte Aktien (Käufe und Verkäufe) in Stück _____	32
Abbildung 2.19: Gehandelte Aktien (Käufe) in Stück _____	33
Abbildung 2.20: Gehandelte Aktien (Verkäufe) in Stück _____	33
Abbildung 2.21: Gehandelte Aktien in Stück und Vier-Tages-Intervall _____	34
Tabelle 2.7: Ergebnisse des APSM97 verglichen mit tatsächlichem Endergebnis _____	36
Abbildung 2.22: Zeitlicher Verlauf der Summe der Kurse aller Aktien _____	37
Abbildung 2.23: Kursverlauf und Handelsvolumen im ÖVP-Markt _____	38
Abbildung 2.24: Gehandelte Aktien (Käufe und Verkäufe) – ÖVP _____	38
Abbildung 2.25: Kursverlauf und Handelsvolumen im SPÖ-Markt _____	39
Abbildung 2.26: Gehandelte Aktien (Käufe und Verkäufe) – SPÖ _____	40
Abbildung 2.27: Kursverlauf und Handelsvolumen im FPÖ-Markt _____	41
Abbildung 2.28: Gehandelte Aktien (Käufe und Verkäufe) – FPÖ _____	41
Abbildung 2.29: Kursverlauf und Handelsvolumen im Grüne-Markt _____	42
Abbildung 2.30: Gehandelte Aktien (Käufe und Verkäufe) – Grüne _____	43
Abbildung 2.31: Kursverlauf und Handelsvolumen im LiF-Markt _____	44
Abbildung 2.32: Gehandelte Aktien (Käufe und Verkäufe) – LiF _____	44
Abbildung 3.1: In Forschungspraxis bewährte Methoden der Umfrageforschung _____	49
Abbildung 3.2: Normalverteilung _____	53
Abbildung 3.3: Elemente der Strukturierung eines Objektbereichs _____	56
Tabelle 3.1: Beispiel einer Typologie, die Formen des Verhaltens von Merton _____	59
Abbildung 3.4: Grundgesamtheit und Stichprobeneinheit _____	60
Tabelle 3.2: Qualität des Adressenmaterials für österreichische Umfragen _____	69
Tabelle 3.3: Ausschöpfungsraten nationaler Großumfr. in Europa u. Nordamerika _____	71
Abbildung 3.5: Darstellung des True-Score-Meßmodells _____	75
Tabelle 3.4: Vergleich des Meßqualitätsniveaus zw. Österreich und den USA _____	76
Abbildung 3.6: Das Interview als sozialer Prozeß _____	83
Tabelle 3.5: Länge der Frageneinleitungen, Stimuli, Anzahl der Antwortkategorien _____	93
Tabelle 3.6: Im Juli durchgeführte Befragungen von IFES und SPECTRA _____	94

<i>Tabelle 3.7a: Beob. Häufigkeiten bei den Umfragen von SPECTRA im Juli 97</i>	95
<i>Tabelle 3.7b: Erwartete Häufigkeiten</i>	95
<i>Tabelle 3.8: Anfang Sept. durchgeführte Befragungen von IFES und SPECTRA</i>	96
<i>Tabelle 3.9: Beob. Häufigkeiten bei Umfragen von SPECTRA und IFES Sept. 97</i>	97
<i>Tabelle 3.10: Erwartete Häufigkeiten</i>	97
<i>Tabelle 3.11: Letzte Befragungen von IFES und SPECTRA</i>	97
<i>Tabelle 3.12: Beob. SPÖ-Häufigkeiten bei den letzten beiden IFES-Umfragen</i>	98
<i>Tabelle 3.13: Erwartete SPÖ-Häufigkeiten</i>	98
<i>Tabelle 3.14: Beob. ÖVP-Häufigkeiten bei den letzten SPECTRA-Umfragen</i>	98
<i>Tabelle 3.15: Erwartete ÖVP-Häufigkeiten</i>	98
<i>Tabelle 3.16: Die letzten Ergebnisse verglichen mit tatsächlichen Wahlergebnis</i>	99
<i>Tabelle 3.17: Positive Meinung von Spitzenkandidaten bei IFES und SPECTRA</i>	100
<i>Tabelle 3.18: Besten Eindruck bzw. Landeshauptmannpräferenz</i>	100
<i>Abbildung 3.7: Lambach-Frage</i>	101
<i>Abbildung 3.8: Welche Partei hat bez. Lambach gleiche Position wie sie selbst?</i>	101
<i>Abbildung 3.9: Und welcher Politiker ist für diese Affäre hauptverantwortlich?</i>	102
<i>Abbildung 4.1: Experimente in den Wirtschaftswissenschaften</i>	103
<i>Tabelle 4.1: Vergleich der Ergebnisse von APSM und letzten Meinungsumfragen</i>	106
<i>Tabelle 4.2: APSM versus Meinungsforschung – Bundespräsidentenwahl 1998</i>	107
<i>Abbildung 4.2: Kursverläufe und Handelsvolumen beim BP-Wahlmarkt</i>	107
<i>Tabelle 4.3: MAE und RSME bei den verschiedenen Projekten</i>	108
<i>Abbildung 4.3: Mean Average Error bei APSM und Umfragen</i>	109
<i>Abbildung 4.4: Root Square Mean Error bei APSM und Umfragen</i>	109
<i>Abbildung 4.5: Auszahlungsregel im Verzug-Markt</i>	113
<i>Abbildung 4.6: Preisentwicklung und Handelsvolumen im Markt „Verzug“</i>	114

## VII. Literatur- und Quellenverzeichnis

Berg, J., Forsythe, R., Rietz, T.: What Makes Markets Predict Well? Evidence from the Iowa Election Markets, College of Business Administration, University of Iowa, forthcoming in Essay in Honor of Reinhard Selten, Springer Verlag 1996

Buchinger, R.: Landtagswahl, Gemeinderatswahl, Bürgermeisterwahl: Ergebnisse und Analysen, Thesenpapier, SPÖ-Klub, Linz 1997

Davis, D.D., Holt, C.H.: Experimental Economics, Princeton University Press, New Jersey 1993

Forsythe R., Forrest, N., Neumann, G., Wright, J.: The Iowa Presidential Stock Market: A Field Experiment, Research in Experimental Economics, Volume 4, 1-43, 1991

Friedrichs, J.: Methoden empirischer Sozialforschung, 14.Auflage, Westdeutscher Verlag GmbH, Opladen 1980

Glosten, L.R.: Is the Electronic Open Limit Order Book Inevitable?, in: The Journal of Finance, Vol. XLIX, No.4, September 1994

Holm, K.: Die Befragung 1: Der Fragebogen - Die Stichprobe, 2.Auflage, Franke Verlag GmbH, München 1982

Holm, K.: Die Befragung 2: Datenaufbereitung, Tabellenanalyse, Korrelationsmatrix, Franke Verlag GmbH, München 1975

Holm, K.: Die Befragung 3: Die Faktorenanalyse, Franke Verlag GmbH, München 1976



Holm, K.: Die Befragung 4: Skalierungsverfahren, Panelanalyse, Franke Verlag GmbH, München 1976

IFES: Die politische Stimmung in Oberösterreich im Frühsommer 1997, Wien, Juli 1997

IFES: Die politische Stimmung in Oberösterreich 1 Monat vor der Landtagswahl, Wien, September 1997

IFES: Die politische Stimmung in Oberösterreich 1 Woche vor der Landtagswahl, Wien, September 1997

Karmasin, F., Karmasin, H.: Einführung in Methoden und Probleme der Umfrageforschung, Hermann Böhlau Nachf. Ges.m.b.H, Graz 1977

Költringer, R.: Die Interviewer in der Markt- und Meinungsforschung, Service Fachverlag an der WU Wien, Wien 1992

Költringer, R.: Gültigkeit von Umfragedaten, Böhlau, Wien 1993

Költringer, R.: Richtig Fragen heißt besser Messen, in: Schweiger, G. (Herausgeber), Empirische Marketingforschung, Band 10, Mannheim 1997

Michenthaler, G., Projektleiter bei der IFES-Studie zur Oberösterreichischer Landtagswahl, Interview am 28.10.1997

Modigliani, F., Miller, M.H.: The Cost of Capital, Corporation Finance and the Theory of Investment, in: American Economic Review, Vol.48, 261-297, 1958;

Neuwirth, E.: Statistische Methoden für Wahlanalysen und Wahlprognosen, Internet (<http://sunsite.univie.ac.at/Austria/elections/wahlstat.html>), Stand 1997

Ortner, G., Stepan, A.: Political Stock Market Experimente, Working Paper, TU Wien, Internet (<http://ebweb.tuwien.ac.at/apsm/paper1.html>), Juli 1994

Ortner, G.: Experimentelle Aktienmärkte als Prognoseinstrument – Qualitätskriterien der Informationsverarbeitung am Bsp. PSM, Sozial- und Wirtschaftswissenschaftliche Fakultät UNI Wien, 1996

Ortner, G.: Forecasting Markets – An Industrial Application, Part 1 and 2, TU Wien, 1997/1998

Ortner, G.: Firmeninterner Aktienmarkt als Prognoseinstrument, TU Wien, 1997

Ortner, G.: Austrian Electronic Markets, AEM Homepage, TU Wien, Internet (<http://ebweb.tuwien.ac.at/apsm/aem97.html>), Stand 1997

Pawek, S.: The APSM 95/96 – Performance eines politischen Aktienmarktes, Diplomarbeit, TU Wien, 1996

PSE aktuell, Artikel in Siemens-Zeitschrift, 5/97

Smith, V.L.: Experimental Economics, Elgar, Aldershot/Hants 1990

SPECTRA: Countdown zur OÖ-LTW 97, 5.Welle, Linz 1997

Statistischer Dienst des Landes OÖ: Landtagswahl 1997 – vorläufiges Endergebnis, Linz 1997

Watzl, E.: Nachlese zur Landtagswahl, Gemeinderatswahl und Bgm-Direktwahl 1997, Thesenpapier, ÖVP-Klub, Linz 1997