

Landsfader eller landssviker?

En kvantitativ analyse av hvilke faktorer som predikerer grad av storhet blant norske statsministere i perioden 1905-1997

Jon Dagsland Holgersen
masteroppgave



Institutt for sammenliknende politikk
Universitetet i Bergen
Mai 2008

ABSTRACT

In the US there are numerous studies on presidential greatness. It started in 1948 when the historian Arthur M. Schlesinger asked 55 of the leading American historians to rate all the presidents, starting from George Washington, on a scale from 1 to 5. The result was published in Life Magazine in 1949 and became heavily debated. Since then there have been undertaken many analyses concerning which factors predict presidential greatness. Years in the presidency, war years and presidential scandals are generally seen as significant predictors on presidential greatness in the US. Which variables predict greatness among 25 prime ministers in Norway between 1905-1997? Such quantitative analyses have never been undertaken, and therefore this study focuses on a wide approach using multiple regression and OLS method. The data are mainly collected from the Norwegian Bureau of Statistics and Norsk samfunnsvitenskapelig datatjeneste. Of the 21 included independent variables, divided into groups of *institutional*, *socioeconomic* and *biographical* factors, *months as prime ministers* and *unemployment* were found to be significant variables in predicting greatness. These factors account for 53 percent of the variation in greatness. *Months as prime ministers* was the most important factor in predicting greatness. By studying the residuals it was revealed that some prime ministers did not fit into the model. Christian Michelsen, who led the country when it came independent from Sweden in 1905, was heavily underpredicted, probably because the respondents emphasized this event when rating him great.

FORORD

Jeg har mange å takke for at denne masteroppgaven ble fullført.

Først og fremst vil jeg rette en stor takk til min veileder, Gunnar Grendstad, for å introdusere meg for temaet, og for å ha vært behjelpelig med konstruktive innspill gjennom hele perioden.

Takk til Kjell Cordtsen for respondentenes vurderinger og kommentarer, samt bakgrunnsmateriale for artikkelen i 1997. Jeg må også takke Ola H. Grytten på Norges Handelshøyskole for data på arbeidsledige personer før 1960.

Takk til deltakerne i metodegruppen for verdifulle kommentarer på hypotese- og analysekapittel. Jeg vil spesielt takke Eirik Vestrheim og Eirik Meling for deres arbeid med datasettet.

Jeg vil også takke mine medstudenter, Frode og Anders, for konstruktive kommentarer på oppgaven.

Mette og Svein fortjener også takk for gjennomlesning og engasjement gjennom denne prosessen.

Til slutt vil jeg takke Maren for hennes oppmuntrende kommentarer og gode innspill.

INNHOLDSFORTEGNELSE

INNHOLDSFORTEGNELSE	III
TABELLER OG FIGURER	V
1. INTRODUKSJON TIL FORSKNINGSSPØRSMÅLET	1
1.1 STRUKTUR	2
2. PREDIKSJON AV STORHET	4
2.1 HVORDAN GÅ FREM?	4
2.2 OVER- OG UNDERPREDIKERING	5
2.3 BETYDNINGEN AV KONTEKST. USA – NORGE: EN LANG REISE	6
2.4 TIDEN LEGER ALLE SÅR?	8
2.4.1 <i>Hvorfor endres respondentenes oppfatninger?</i>	8
2.4.2 <i>Betydningen av tid i historisk kontekst</i>	9
2.5 OPPSUMMERING	12
3. HYPOTESEKAPITTEL	14
3.1 AVHENGIG VARIABEL	14
3.2 PREDIKSJONSVARIABLER	14
3.2.1 <i>Institusjonelle variabler</i>	17
3.2.2 <i>Sosioøkonomiske variabler</i>	21
3.2.3 <i>Biografiske variabler</i>	22
3.3 OPPSUMMERING AV HYPOTESEKAPITTEL	23
4. DATA, OPERASJONALISERING OG METODE	24
4.1 CORDTSEN-STUDIEN	24
4.2 OBJEKTIVE RESPONDENTER?	27
4.3 OPERASJONALISERING AV VARIABLER	30
4.3.1 <i>Institusjonelle faktorer</i>	30
4.3.2 <i>Sosioøkonomiske faktorer</i>	32
4.3.3 <i>Biografiske faktorer</i>	33
4.4 KVANTITATIV METODE	35
4.5 MULTIVARIAT REGRESJON	35
4.5.1 <i>Mål og tester i regresjon</i>	35
4.5.2 <i>Prediksjon versus kausalitet</i>	39
4.6 MÅLE- OG SPESIFIKASJONSFEIL	41
4.7 KRITERIER I ANALYSEPROSESSEN	42
4.8 UTFORDRINGER KNYTTET TIL TID OG OLS-REGRESJON	44
4.9 OPPSUMMERING AV DATA, OPERASJONALISERING OG METODE	45
5. ANALYSE OG RESULTATER	48

5.1	STRUKTUR	48
5.2	VARIABLENES FORDELING.....	48
5.3	REGRESJONSANALYSER	50
5.3.1	<i>Bivariat regresjon</i>	51
5.3.2	<i>Institusjonelle faktorer</i>	52
5.3.3	<i>Sosioøkonomiske effekter</i>	56
5.3.4	<i>Biografiske effekter</i>	58
5.3.5	<i>Samlede effekter</i>	60
5.4	STABILITET – ANALYSER AV RESIDUALER.....	64
5.4.1	<i>Hvorfor avviker noen statsministere?</i>	67
5.5	PREDIKSJON AV NYE STATSMINISTERE	71
5.6	KONKLUSJON AV EMPIRISKE FUNN	71
6.	KONKLUSJON	75
6.1	OPPSUMMERING AV FUNN	76
6.2	HVOR GÅR VEIEN VIDERE?	78
7.	LITTERATURLISTE.....	80
8.	APPENDIKS	83
	APPENDIKS A – VARIABLEOVERSIKT	83
	APPENDIKS B – RESPONDENTENE I CORDTSEN-STUDIEN.....	85
	APPENDIKS C – KORRELASJONSMATRISE	86
	APPENDIKS D – KRONOLOGISK RANGERING AV NORGES STATSMINISTERE I PERIODEN 1905-1997 .	87
	APPENDIKS E – TEST AV FORUTSETNINGER	88

TABELLER OG FIGURER

Tabell 1 – Statsministerne og deres storhetsscore.....	26
Tabell 2 – Deskriptiv statistikk av variablene i analysen	49
Tabell 3 – bivariat regresjon.....	51
Tabell 4 – institusjonell modell.....	52
Tabell 5 – sosioøkonomisk faktorer	56
Tabell 6 – biografiske faktorer	58
Tabell 7 – endelig modell.....	61
Tabell 8 – Storhetsscore og variabelverdier.....	64
Tabell 9 – Hvem faller utenfor modellen?.....	66
Figur 1 – Variabelgrupper og enkeltvariablers antatte prediksjon på storhetsgrad.....	16

"Presidential 'success', as Bert Rockman has aptly observed, 'is both multidimensional and frequently inextricable from dumb luck'" (Bond, Fleisher & Krutz 1996: 106).

"If number one can be determined in athletics, in rental car agencies, and in fast-food, why not in the presidency?" (Murray & Blessing 1983: 535).

1. INTRODUKSJON TIL FORSKNINGSSPØRSMÅLET

Som leder for regjeringen er statsministeren satt til å styre landet på en best mulig måte. Derfor er statsministerrollen kanskje den viktigste jobben i Norge. I perioden 1905-1997 har landet hatt 25 ulike statsministere. Hvorfor blir noen vurdert som gode statsministere mens andre betegnes som dårlige? Problemstillingen denne analysen skal prøve å besvare er: *Hvilke faktorer predikerer graden av storhet blant norske statsministere i perioden 1905-1997?* Dette er første gang en slik analyse gjennomføres i Norge, og oppgaven vil således bidra til ny kunnskap på det statsvitenskapelige feltet.

Analysen tar utgangspunkt i historikere og samfunnsviters vurderinger av statsministerprestasjonene til norske regjeringsledere. Akademikerne har gjennom sin undervisning og som samfunnsdebattanter anledning til å påvirke både studenters og den øvrige befolkningens oppfatninger av våre statsministere (se Murray & Blessing 1983: 537). Hvilke kriterier som legges til grunn når de vurderer landets regjeringsledere vil derfor kunne være av interesse både for akademikere, media, politikere og den allmenne befolkning.

I USA er det lang tradisjon for at historikere og samfunnsvitere vurderer prestasjonene til amerikanske presidenter (Simonton 2001: 293). Første gang dette ble gjort på en systematisk måte var i 1948 da 55 ledende amerikanske historikere fikk oppdraget fra historikeren og Harvard-professoren Arthur M. Schlesinger. Respondentene fikk selv bestemme hvilke kriterier de ville legge til grunn for sine vurderinger. Resultatet ble publisert i Life Magazine i 1949. Artikkelen vakte stor interesse og førte til mye diskusjon (Schlesinger 1979). I 1962 gjentok Schlesinger prosedyren på oppdrag fra The New York Times Magazine. Sønnen, Arthur M. Schlesinger jr., videreførte sin fars arbeid da han gjorde en tilsvarende studie på oppdrag fra samme magasin i 1996. Siden oppstarten har forskningsfeltet utviklet seg til å omfatte mer sofistikerte analyser av presidenters storhetsgrad. Et eksempel på dette er at respondenter vurderer presidentenes prestasjoner innenfor bestemte politiske områder¹ (se for eksempel Lønnstrøm & Kelly 2003: 626).

¹ Politisk lederskap, utenrikspolitikk, nasjonal politikk, personlige egenskaper og historisk påvirkning er noen slike definerte kategorier (Schlesinger 1997:183).

I 1997 fikk Verdens Gang-journalisten Kjell Cordtsen 20 av Norges fremste forskere innen historie og statsvitenskap, blant dem tre politiske redaktører, til å vurdere statsministerprestasjonene til samtlige norske regjeringsledere siden 1905 (Cordtsen 1997b). Cordtsen bygget studien sin på Schlesingers metode, og respondentenes vurderinger er således utgangspunkt for verdiene på den avhengige variabelen i denne studien. Cordtsen-undersøkelsen resulterte i at Einar Gerhardsen ble vurdert som den største norske statsminister gjennom tidene mens Christian Michelsen kom på andre plass. Nederst på listen endte Jens Hundseid med Thorbjørn Jagland på nest siste. Det at sistnevnte i den aktuelle artikkelen ble sammenlignet med Jens Hunseid fikk Thorbjørn Berntsen til å komme med følgende reaksjon i Verdens Gang:

”I dette ’råttekjøret’ mot Jagland var det vel bare dette som manglet, nemlig en dårlig skjult antydning om at Jagland nærmest er å sammenligne med Hunseid som hadde Quisling som regjeringsmedlem. Dere nekter dere sannelig ikke noe! ’KGB-agenten’ Jagland er jo lansert fra før” (Berntsen 1997).

Meningene er delte når historikere og statsvitere skal vurdere statsministeres prestasjoner. Per Borten, som regjerte mellom 1965-1971, ble for eksempel omtalt som både betydelig og en fiasko av ulike respondenter (Cordtsen 1997a). Dette tyder på at de legger forskjellige vurderingskriterier til grunn for sine bedømmelser. I denne sammenheng vil det være interessant å identifisere eventuelle fellesnevnerne i respondentenes vurderingene.

1.1 Struktur

Det neste kapittelet fokuserer tendenser fra den internasjonale litteraturen, som hovedsakelig er amerikansk. På grunnlag av dette vil det diskuteres hvilke konsekvenser kontekstuelle ulikheter mellom først og fremst Norge og USA har for denne type studier. Problemstillinger knyttet til betydningen av tid i analyser av regjeringsledere vil også vektlegges. I denne sammenheng fokuseres det først på årsaker til hvorfor noen presidenter blir vurdert høyere/lavere med tiden. Deretter blir betydningen av tid diskutert i forhold til de ulike kontekstuelle forutsetningene hver statsminister regjerte under.

I hypotesekapittelet blir variablene i denne analysen presentert og drøftet. Det blir her fokusert på bakgrunnen og inspirasjonen for de inkluderte faktorene, samt deres antatte prediksjon på storhetsgrad.

Kapittelet *data, operasjonalisering og metode* består av tre hoveddeler. Den første tar for seg datagrunnlaget i Cordtsen-studien, og mulige svakheter knyttet til respondentenes objektivitet. Den andre hoveddelen fokuserer på operasjonaliseringen av variablene, mens den tredje tar for seg metodiske problemstillinger knyttet til analysen og de verktøyene som vil bli brukt.

Analysen er delt inn i to sekvenser. Den første predikerer storhet ved hjelp av regresjonsanalyse. Den endelige prediksjonsmodellen vil inneholde de variablene som er sentrale for prediksjon av statsministerprestasjoner. I analysens andre sekvens vil det ved å se på tendenser i residualene, forskjellen mellom faktiske og predikerte verdier, fokuseres på hvor godt ulike statsministere blir predikert av modellen. Her vil det også bli gjennomgått eksterne faktorer som kan være årsakene til mulig avvikende verdier hos statsministere. I den amerikanske litteraturen er det gjort funn som viser at det er betydelig variasjon i residualverdiene mellom USAs presidenter (Simonton 1986: 272). Deretter blir statsministere, som påvirker modellen og de enkelte uavhengige variablene uforholdsmessig mye, diskutert. Basert på den endelige prediksjonsmodellen vil analysen også predikere storhetsgraden til Norges statsministere etter 1997.

2. PREDIKSJON AV STORHET

En utfordring med denne studien er at det foreligger lite konkret teori på statsministerprestasjoner i Norge. I USA har man kommet lengre på dette området fordi det er mer kunnskap om hvilke variabler som predikerer presidentenes storhetsgrad. Kapitlet retter fokus mot den internasjonale litteraturen og utfordringer knyttet til bruken av variabler som hovedsakelig er tilpasset den amerikanske konteksten. I tillegg vil betydningen av ulike aspekter av tid bli diskutert i forhold til påvirkning på respondentenes vurderinger.

Å definere storhet er ikke uproblematisk. Winter mener at presidenters storhet kan være umulig å definere på en objektiv måte fordi "we can never know all the facts about a president's actions and what independent effects these actions had on historical outcomes" (Winter 1987: 199). Dette er et viktig poeng som også aktualiserer utfordringen ved å studere et fenomen det er mangelfull forskning på. Likevel er Cordtsen-undersøkelsen, som vil bli gjennomgått i kapittel 4, i samsvar med måten forskningsfeltet ble introdusert på i USA. Med dette som utgangspunkt kan det argumenteres for at fremgangsmåte er passende.

Resultatene fra Cordtsen-studien har gjort det mulig å foreta kvantitative analyser av hvilke variabler som kan predikere graden av storhet blant norske statsministere, samt vise hvem som blir godt og dårlig predikert innenfor modellens rammer. I denne analysen er prediksjon viktigere enn kausalitet, og årsaker til dette vil bli diskutert i kapittel 4.

2.1 Hvordan gå frem?

Den manglende forskningen i Norge gjør denne studien interessant, men også utfordrende. Med lite teori på området er det vanskelig å vurdere hvilke variabler og data som bør benyttes. Dette kan føre til at variabler som senere viser seg å være irrelevante blir inkludert og visa versa. Det er imidlertid vanskelig å unngå disse utfordringene når man foretar en studie som i stor grad er eksplorerende.

Det foreligger noe litteratur på statsministerprestasjoner i Norge, hovedsakelig av historikere og biografiforfattere. Hypoteser som fremsettes i denne studien er likevel aldri tidligere testet kvantitativt. Et eksempel på en variabel som er hentet fra historikeres generelle beskrivelser av statsministernes prestasjoner, er hvorvidt statsministeren har vært leder for koalisjonsregjering. Enkelte andre variablene som kan predikere storhet vil bli trukket ut fra

biografilitteraturen. Den kanskje viktigste inspirasjonskilden for inkluderte variabelgrupper og enkeltfaktorer er likevel internasjonal, og da hovedsakelig amerikansk, forskning. Årsaker til hvorfor statsministere ble innsatt og avsatt er to eksempler på variabelgrupper fra studier i USA. Noen faktorer er derimot inkludert på et mer intuitivt grunnlag. Disse nevnes ikke eksplisitt i litteraturen, men det kan tenkes at de likevel har en prediksjonseffekt. Dette gjelder blant annet Stein Rokkans beskrivelse av det kraftige skillet i norsk politikk i 1927 som gir så store endringer at vi kan snakke om en ny politisk periode i landet.

2.2 Over- og underpredikering

I boken *Christian Michelsen – politikeren* antydes det at statsministeren var en middelmådig politiker, men at han var på rett plass til rett tid (Wyller 1975: 204-206 og 212-216). Det argumenteres således for at det var de kontekstuelle omstendighetene med nasjonal uavhengighet som førte til Michelsens høye status i historien. Kjølås deler dette synet, og det styrker antakelsen om at løsrivelsen fra Sverige i 1905 kan ha effekt på vurderingen av Michelsen (Kjølås 1999: 83).

Dean Keith Simonton er en produktiv forsker innenfor prediksjon av storhet blant amerikanske presidenter. Ifølge hans modell er George Washington, John Harding og John F. Kennedy alle underpredikert i forhold til grad av storhet. Dette innebærer at respondentene vurderer disse presidentenes prestasjoner høyere enn modellen predikerer (1986: 272). For eksempel er statusen til George Washington, som var amerikansk president i perioden 1789-1797, ofte knyttet til rollen som "[t]he father of our country and as the precedent-setting president" (Simonton 1986: 272).

Betegnelsene *landsfader* og *den første president* er interessant i et komparativt perspektiv. Både Michelsen og Gerhardsen har blitt tillagt rollen som Norges landsfedre, og det forventes derfor at de blir underpredikert i analysen. Historiker og forfatter Rune Slagstad bruker i *De nasjonale strateger* uttrykket *landsfader* som inngang til å beskrive Gerhardsens regjeringstid (1998: 355). Det er også tittelen på Egil Helles bok om samme mann (Helle:1987). Michelsen kan dessuten sammenlignes med Washington i forhold til rollen som den første statsminister i det selvstendige Norge. Derfor blir han trolig mer underpedikert enn Gerhardsen.

Andre statsministere blir trolig også over- eller underpredikert. Dette vil avhenge av hvorvidt prediksjonsvariabler som ikke blir tatt med i analysen har påvirket respondentenes bedømmelser av statsministerne. Hva disse faktorene kan være, og hvordan de påvirker vurderingen, blir diskutert i analysekapittelet fordi man her har konkrete tall på prediksjonen av de ulike statsministernes prestasjoner.

2.3 Betydningen av kontekst. USA – Norge: en lang reise

En del av prediksjonsvariablene i den internasjonale litteraturen er enkle å overføre til norsk kontekst mens andre er mer problematiske, og også mindre relevante, for denne studien. Den kontekstuelle variabelproblematikken aktualiserer det som Sartori kaller *conceptual travelling* og *conceptual stretching* (sitert i Collier & Mahon 1993). Tilpasningen som gjøres for å inkludere nye case i eksisterende modeller, og gjøre begreper gjeldende for disse, *conceptual travelling*, kan føre til mangelfullt sammenligningsgrunnlag når begrepene ikke passer de nye casene, *conceptual stretching* (Collier & Mahon 1993: 845). Teorien er ofte brukt for å belyse problematikken som oppstår når man skal kategorisere og lage fellesbetegnelser for ulike typer regimer i verden. I denne sammenheng kan det diskuteres om man kan operere med universelle kategorier, som for eksempel *demokrati* og *diktatur*, når kontekstene statene opererer innenfor er ulike.

I denne studien brukes flere variabler som i utgangspunktet er tilpasset analyser av amerikanske presidenter. Selv om resultatene i Norge og USA i liten grad sammenlignes eksplisitt er det viktig å være seg bevisst problematikken. Eksempler på faktorer som kan knyttes til begrepene *conceptual travelling*- og *stretching* blir gjennomgått i variabeldiskusjonen i neste kapittel. utfordringer i forhold til kodingen av disse variablene vil diskuteres i kapittel 4. Overføring og omkoding av variabler trenger ikke å være et problem i denne studien, men slike endringer kan i mange tilfeller gjøre en sammenlikning mindre verdifull fordi man strekker begrepene og ender opp med målinger som bygger på ulike forutsetninger.

Andre utfordringer knyttet til *conceptual travelling*- og *stretching* omfatter institusjonelle ulikheter mellom USA og Norge, blant annet presidentialisme versus parlamentarisme og toparti- versus flerpartisystem. Linz mener parlamentarismen begrenser regjeringslederens makt, spesielt i tilfeller med koalisjons- og mindretallsregjeringer. Under slike omstendigheter er ”prime ministers much closer to being on an equal footing with their

fellow ministers than presidents will ever be with their cabinet appointees” (Linz 1990: 62). På grunn av at presidenter vanligvis har en mer dominerende maktposisjon enn statsministere er det lettere å holde presidenten ansvarlig for sine handlinger, og dermed måle deres prestasjoner. Parlamentarismen krever også at regjeringens autoritet er fullstendig avhengig av tillit i den lovgivende forsamling. Dette gjør at statsministeren ikke kan handle egenrådig, og fører til at regjeringslederen får en mindre dominerende stilling, noe som igjen svekker grunnlaget for å vurdere deres prestasjoner.

Topartisystemet i USA åpner opp for mer sentrert presidentmakt enn det flerpartisystemet i Norge gir statsministeren. Dette har nær sammenheng med valgsystemet Lijphart (1999: 12,53 og 67) og gjør at presidenten vanligvis fullfører sin regjeringsperiode. I Norge er dette annerledes. Mellom 1905 og 1945 var det for eksempel bare statsministeren Gunnar Knudsen som fullførte sin periode². Selv om presidentene ikke vil kunne operere helt på egenhånd har de større handlefrihet enn norske statsministere. Dette gjør at presidenten i større grad kan holdes direkte ansvarlig for politikken som føres i en regjeringsperiode.

Disse grunnleggende institusjonelle ulikhetene blir spesielt viktige å ta hensyn til i eventuelle kommende studier som eksplisitt vil sammenligne storhet mellom norske statsministere og amerikanske presidenter. Forskjellene er også relevante for sammenligninger mellom land som er antatt likere. Westminster-modellen³ innebærer at ett parti oftest har rent flertall i parlamentet, og systemet fører til ulike forutsetninger for vurderinger av statsministeres prestasjoner enn i Norge. For eksempel hevdes det at Westminster-modellen fører til sterkere regjeringsmakt og mer politisk stabilitet enn i typiske flerpartisystemer (diskutert i Heidar & Berntzen 1998: 60 og 127). Da denne analysen trekker paralleller mellom studier fra ulike land og politiske systemer, tas det hensyn til utfordringene knyttet til sammenligning av variabler og storhetsscore. Variablene som omhandler institusjonelle ulikheter blir diskutert i neste kapittel samt i operasjonaliseringen av variablene.

² For statsministere og regjeringsperioder se: <http://samfunnsveven.no/TT/U40.cfm>

³ Systemet har sin opprinnelse i Storbritannia og idealformen innebærer et topartisystem med sentralisert statsmakt der det partiet som vinner parlamentsflertallet får enerett på regjeringsmakten (Heidar & Berntzen 1998: 60).

2.4 Tiden leger alle sår?

Ulike aspekter av tid og historisk kontekst har trolig betydning for respondenters vurdering av statsministerprestasjoner. Her vil tid først knyttes til endringer i respondentenes bedømmelser av regjeringslederens prestasjoner. Deretter vil fokuset rettes mot tid i forhold til de ulike kontekstuelle forutsetningene hver statsministers regjerte under.

2.4.1 *Hvorfor endres respondentenes oppfatninger?*

I den amerikanske litteraturen har det blitt argumentert for at noen presidenter systematisk blir vurdert høyere/lavere etter hvert som tiden går. Et eksempel på førstnevnte er Dwight D. Eisenhower som har gått fra 22. plass i Schlesinger 1962-studien til niende i to ulike studier fra henholdsvis 1982 og 1996 (Schlesinger 1997:182). Lignende ble Ronald Reagan rangert som nummer 25 av 39 i Schlesinger 1996-studien mens han endte på åttende plass i en ny og mer konservativ undersøkelse i 2000 (Taranto 2004: 1-3). Det blir nå fokusert på fire mulige årsaker til slike utviklinger.

En grunn til endringer i respondentenes vurderinger kan være at presidenters beslutninger alltid virker mer kontroversielle når de skjer, og at det derfor er lettere å dømme disse negativt på kort sikt (Schlesinger jr. 1997: 183). Etter hvert som tiden går, og handlingene blir sett i et større perspektiv, kan derimot respondentenes vurderinger endre seg. I USA brukes gjerne krig som eksempel på et fenomen som virker brutalt når det forekommer, men der respondentene og den øvrige befolkningens holdninger blir mer positiv med tiden. "War seems to reduce presidential popularity as measured by contemporary polls but to produce an image of greatness in historical perspective" (Nice 1984: 453). I norsk sammenheng kan det også vise seg at kontroversielle beslutninger blir vurdert mer positivt med tiden. Dette vil bli diskutert i analysekapittelet.

En annen årsak til endret oppfatning av regjeringslederens prestasjoner er informasjon om presidenters handlinger, som først blir tilgjengelige etter at personen har gått av (Bose & Landis 2003: 89). "Changes in a presidents ranking well after his term are likely to be due to new revelations about the president or his administration and/or to new interpretations" (Lønnstrøm & Kelly 2003: 631). Et eksempel på dette er en artikkel som ble publisert i Washington Post 30.04.08. (Pincus 2008). Her omtales nylig frigitte dokumenter fra George Washington University's nasjonale sikkerhetsarkiv som viser at den tidligere amerikanske presidenten Dwight D. Eisenhower i 1958 trosset råd fra generaler i eget luftforsvar og

forhindret bruk av atombomber mot Kina⁴. Selv om man ikke med sikkerhet kan si at frigivelsen av et slikt dokument vil føre til mer positiv vurdering av Eisenhower, er sjansene store for at denne informasjon vil sette ham i et bedre lys enn tidligere.

En tredje årsak til endringer i respondentenes vurderinger går på holdningsendringer i samfunnet. Dette er tilfellet med Andrew Johnson som var amerikansk president i perioden 1865-1869. Som tilhenger av hvite makthavere blir han stadig vurdert lavere av respondenter i kåringer av store presidenter (Schlesinger jr. 1997: 185). Dette har sammenheng med samfunnsendringene som har skjedd siden Johnson regjerte.

Theakston og Gill har studert statsministerprestasjoner i Storbritannia. De mener at selv svært upopulære regjeringsledere kan få endret sin status med tiden, noe statsministeren Lloyd George er et eksempel på (2006: 210). Han var svært omstridt etter å ha vært innblandet i en korrupsjonsskandale mens han regjerte rundt 1920, men er nå rangert blant de fem største britiske statsministre uten at Theakston og Gill har klare formening om årsaken til dette.

På tross av at det her er presentert argumenter som sier at tid har betydning for endringer i vurderinger av regjeringsledere, er det uenighet i det akademiske miljøet i USA om hvorvidt dette stemmer. Simonton mener at det ikke er noen tendens til at respondentenes bedømmelser endrer seg systematisk over tid (1986: 272). Lindstrøm og Kelly støtter også opp om dette (2003: 633-634). Problemet med å måle slike effekter i denne studien er at det ikke er foretatt lignende analyser i Norge. Derfor kan man ikke se om vurderingene av en statsminister endrer seg i positiv eller negativ retning over tid. Likevel er det relevante problemstillinger i videreføringen av forskningsfeltet fordi de gjør oss mer kritiske til prediksjonsvariablene.

2.4.2 Betydningen av tid i historisk kontekst

Ettersom ingen norske statsministre har regjert under helt like forhold kan det diskuteres hvorvidt det er mulig å predikere storhetsgraden deres. Tim H. Blessing har studert presidentprestasjoner i USA. Han mener at "the factors which emerge as being, by wide margins, most determinate, are those factors related to a presidency's place in time" (1995:

⁴ Bakgrunnen for diskusjonen om bruk av atomvåpen var at Kina var i ferd med å blokkere Taiwanstredet. Etter noen turbulente måneder valgte Kina i oktober 1958 å opphøre all ild, og situasjonen roet seg.

1). Theakston og Gill belyser utfordringen knyttet til sammenligningsgrunnlaget mellom amerikanske presidenter med å argumentere for at “[t]he problem with comparative scoring exercises is that no two leaders are ever dealt the same hand” (2006: 211). I den internasjonale litteraturen har disse utfordringene vært opphav til mye diskusjon. Et sentralt spørsmål i forhold til sammenligningsgrunnlaget er hvorvidt man kan foreta vurderinger av prestasjoner når forhold som statsministeren selv ikke har makt over kan avgjøre deres handlinger og grad av suksess. Internasjonale politiske endringer og økonomiske konjunkturer, samt krig, er eksempler på dette.

Berge Furre argumenterer for at årsaken til Kåre Willoch og Høyres suksess på 1980-tallet var en europeisk høyrebølge med blant annet Margaret Thatcher i Storbritannia (1991: 402-403). Furre mener ”politiske vindar bles til fordel for eit parti som baud skattelette, mindre regulering, mindre offentlig forbruk, mindre stat, meir handlefridom for den initiativrike” (1991: 402). Dette argumentet støttes også i et debattskrift til Kåre Willoch (Brende, Clemet, Norvik & Stray 1993: 207-209). I studien av fire britiske statsministere mener Kevin Theakston at en stor del av årsaken til Margaret Thatchers høye storhetscore blant britiske statsministere er at hun var ”heldig med tidevannet” i betydningen den politiske høyrevridningen i Europa (2003: 105).

Betydningen av kontekstuelle forhold blir også knyttet til Einar Gerhardsen. Finn Olstad beskriver ham som ”[e]n mann av sin tid, en arbeidergutt som ble formet i Oslo i det radikale oppbruddets tid, og at dette var sentrale elementer for hans suksess (1999: 427). Hansson og Teigene sier det enda mer direkte: ”Nå er det en tid for alt – og Einar Gerhardsen var en leder av sin tid” (1992: 323).

En statsministers storhetsgrad vil også være avhengig av forgjengerens makt, innflytelse og valgene han foretok. I *Oscar Torp – en politisk biografi* blir det skrevet at Torp hadde en vanskelig oppgave med å følge etter Einar Gerhardsen, og at det egentlig var sistnevnte som hadde den reelle makten i Arbeiderpartiet i årene 1951-1955 (Lahlum 2007: 283, 289-291 og 356). Dette gjorde at ”odds var imot ham, og han var klar over det selv” (Lahlum 2007: 283). Statsministere må forholde seg til valgene den forrige regjeringssjefen har foretatt, og dette begrenser handlingsrommet og muligheten for å få gjennomslag for egen politikk. Tim H. Blessing sammenligner situasjonen med et spindelweb der ”one president follows another, and many presidents become entrapped in the webs which their predecessors have spun”

(1995: 2 og 6). Likevel har analyser fra USA vist at autokorrelasjon ikke er et vanlig problem i studier av presidentprestasjoner (Simonton 1986).

Ross mener det er et problem at respondenter kun fokuserer på individ og således tillegger presidentene egenskaper de ikke selv har kontroll over. Ved å overse de situasjonsbetingede begrensningene gjør respondentene det Ross kaller "the fundamental attribution error" (sitert i Simonton 1986: 279). Dette kan særlig knyttes til vurderinger av statsministere som var ledere i perioder da det foregikk store endringer. Eksempler på dette i Norge er unionsoppløsningen i 1905 og nedgangsperioden i 1920- og 30-årene. Spørsmålet er hvor store mulighetene egentlig var for å gjøre en god jobb når det var internasjonale nedgangstider, en faktor statsministerne ikke selv kunne påvirke. Lignende kan det diskuteres hvor stor innflytelse Michelsens hadde på løsrivelsen fra Sverige. Ross argumenterer for at hendelser og situasjoner ofte blir tilskrevet ledere uten at personen nødvendigvis hadde noen kausal effekt på resultatet (sitert i Simonton 1986: 279). I slike tilfeller blir ofte situasjonsbetingede hendelser ignorert selv om de er den egentlige årsaken til hendelsen. Leo Tolstoy skrev i *War and Peace* (sitert i Simonton 1986: 279) at "the so-called great men are labels giving names to events, and like labels they have but the smallest connection with the event itself." Denne studien unngår likevel disse utfordringene fordi den analyserer prediksjon på storhet, ikke kausale forhold mellom de uavhengige variablene og den avhengige.

Historisk sett er ofte muligheter for å få gjennomført endringer i politikken svært ulike. For eksempel ville det i etterkrigsperioden vært vanskelig for etterfølgerne til Gerhardsen å endre den oppbyggingen av sosialdemokratiet som da var påbegynt. Franklin D. Roosevelt, som introduserte *The New Deal*⁵ i USA, mente at hvis de påfølgende presidentene ikke fulgte prinsippene han hadde lagt til grunn for reformene så ville man ikke høre mer om denne presidenten eller personens parti igjen (Blessing 1995: 6).

Det kan også være en utfordring at det faktisk ikke skjer spesielle hendelser i en periode, og at man derfor ikke får muligheten til å bli vurdert som en stor leder. Theodore Roosevelt mente at perioden han selv regjerte under aldri ga ham muligheten til å bli en stor president i USA, og at "if Lincoln had lived in times of peace, no one would know his name now"

⁵ Innføring av en serie reformer, hovedsakelig i perioden 1933-38, for å løse den økonomiske krisen som hadde oppstått på grunn av nedkonjunkturer i mellomkrigstiden (Ranson 2005: 768-769).

(sitert i Adler 2003: 6). Når det er sagt fikk likevel Roosevelt Nobels fredspris i 1906. Enkelthendelser kan også påvirke respondentenes vurderinger. I Storbritannia blir for eksempel Suez-krisen⁶ regnet for å være hovedårsaken til at Anthony Eden har blitt rangert lavest i kåringer av Storbritannias statsministere (Theakston & Gill 2006: 198).

På tross av ulike kontekstuelle forutsetninger kan det argumenteres for at statsministere alltid må handle best ut fra de historiske rammene i deres regjeringstid, og at store statsministere vil takle denne utfordringen bedre enn dem som blir vurdert mindre store. Frankrikes tidligere president Charles de Gaulle har uttalt at ”a statesman may be determined and tenacious, but, if he does not understand the character of his time, he will fail” (sitert i Schlesinger jr. 1997: 186). Schlesinger argumenterer i denne sammenheng for at store presidenter har en bedre forståelse av behovene, frykten og drømmene til folket.

Man må også kunne forvente at respondentene tar i betraktning ulike kontekstuelle forhold når de vurderer statsministerprestasjoner (Ballard og Suedfeld sitert i Theakston & Gill 2006: 211). En gjennomgang av Cordtsen-respondentenes utfyllende vurderinger av norske statsministerprestasjoner styrker denne antakelsen (Cordtsen 1997a). Dette øker verdien av denne type studier. Potensielle utfordringer knyttet til respondentenes subjektive påvirkningsmuligheter på statsministeres storhetsgrad vil bli nærmere gjennomgått i kapittel 4.

2.5 Oppsummering

Dette kapitlet har tatt for seg problematiske aspekter i forhold til at denne analysen er første av sitt slag i Norge. Dette gjør at mange variabler er hentet med inspirasjon fra hovedsakelig amerikansk forskning, noe som ikke er uproblematisk. Diskusjonen omkring conceptual travelling- og stretching, og konsekvensene av disse, sto her sentralt. Blant annet innebærer de institusjonelle ulikhetene mellom Norge og USA at presidenter i USA har en mer dominerende posisjon på grunn av større maktsentralisering. Dette gir dem større handlingsrom og gjør at man i større grad kan holde dem ansvarlig for politikken som føres enn hva tilfellet er i Norge. I Norge må statsministeren ofte forholde seg til koalisjons- og mindretallsregjeringer, samt støtte i Stortinget. Konsekvensen av dette er at det trolig er

⁶ Militært oppgjør i perioden 1910-1956 mellom Storbritannia, Frankrike og Israel på den ene siden, og Egypt på den andre. Konflikten endte med at felles militær traktat i 1956 etter press fra USA og Sovjetunionen i FNs sikkerhetsråd (Lund 2000: 260).

lettere og mer presist å predikere storhet blant amerikanske presidenter enn norske statsministere. De kontekstuelle forskjellene gjør også at flere sentrale variabler i USA, som for eksempel krig, ikke er like interessante for denne studien.

Det ble deretter rettet fokus mot betydningen av tid og historisk kontekst for vurderingene av regjeringsledere. Schlesinger jr. er blant dem som mener at respondentene systematisk bedømmer enkelte presidenter høyere og lavere etter hvert som tiden går (1997: 183). Krig er et eksempel på en hendelse som ofte virker dramatisk når den pågår, men i et historisk perspektiv blir beslutningen gjerne vurdert mer positivt. Dessuten blir bakgrunnen for en leders handlinger ofte ikke offentliggjort før etter en viss tid, og slik informasjon kan ha betydning for vurderinger av regjeringslederne. I tillegg kan holdningsendringer i samfunnet føre til at respondentenes bedømmer regjeringsledere mer positivt/negativt med tiden. Det er likevel uenighet om betydningen av tid i prediksjon av storhet. Ifølge Simonton er det ingen klare tendenser til at vurderingene av amerikanske presidenter endrer seg systematisk over tid (1986: 272).

Angående de kontekstuelle ulikhetene mellom hver statsministers regjeringstid har det blitt reist kritikk mot sammenligningsgrunnlaget. Et spørsmål er om man kan predikere grad av storhet når ingen statsminister har samme utgangspunkt (se for eksempel Theakston & Gill 2006). Krig og økonomiske konjunkturer er eksempler på slike faktorer. Man kan også diskutere hvilket kausalt forhold det er mellom hendelser og personer. Christian Michelsens betydning for unionsoppløsningen i 1905 ble her nevnt. Denne delen ble konkludert med at en statsminister må handle best innenfor de gitte kontekstuelle rammene. Dessuten forventes det at respondentene tar hensyn til kontekst, og denne antakelsen ble styrket av Cordtsen-respondentenes uyllende kommentarer (Cordtsen 1997a). Dette øker verdien av denne studien.

Etter denne diskusjonen av inspirasjonskilder for variablene, samt problematiske aspekter knyttet til slike studier, vil de konkrete variablene bli gjennomgått i neste kapittel.

3. HYPOTESEKAPITTEL

I dette kapittelet blir den avhengige variabelen og de uavhengige presentert og diskutert. Disse danner grunnlaget for å predikere storhetsgraden til norske statsministere i perioden 1905-1997.

3.1 Avhengig variabel

Avhengig variabel i denne analysen er graden av *storhet* blant norske statsministere. Storhetsscoren tar utgangspunkt i respondentenes subjektive vurderinger av hvilke faktorer som skal legges til grunn. Dette samsvarer med måten Schlesinger og Schlesinger jr. har utført sine undersøkelser på, samt Theakston og Gills analyse av britiske regjeringsledere (2006: 197). Det har således blitt gitt stor tillit til at respondentene har tilstrekkelige forutsetninger for å bedømme statsministerprestasjoner. Schlesinger jr. forsvarte vurderingsgrunnlaget ved å argumentere for at respondentene hadde god nok faglig kompetanse til å skille mellom store og dårlige presidenter, og at det derfor ikke var nødvendig med mer definerte vurderingskriterier. ”It was supposed that historians would know greatness – or failure – when they saw it, as Justice Potter Stewart famously said of pornography” (Schlesinger jr. 1996: 46). Argumenter for og imot denne fremgangsmåten vil bli gjennomgått nærmere i neste kapittel.

3.2 Prediksjonsvariabler

Ifølge Simonton (1986: 262) er følgende fem variabler sentrale i prediksjon av presidenters storhetsgrad: (1) hvor lenge presidenten var innsatt, (2) hvor lenge nasjonen var i krig i presidentperioden, (3) hvorvidt presidenten overlevde et attentatforsøk, (4) hvorvidt en stor skandale oppsto da personen styrte og (5) hvor mange publikasjoner han hadde i forkant av innsettelsen. Etter at det er kontrollert for korrelasjon, og foretatt multivariate regresjonsanalyser, er det likevel kun *antall år som president*, *antall krigsår* i perioden og *presidentskandaler* som gjennomgående fungerer som sikre prediktatorer på storhet (Simonton 1986: 268 og 273).

Av disse faktorene vil denne studien kun benytte seg av lengden på statsministernes innsettelsesperiode. Dette belyser de kontekstuelle ulikhetene mellom USA og Norge diskutert i forrige kapittel. For eksempel blir brutto nasjonalprodukt (BNP) regnet for å være av mindre betydning enn krig i vurderingen av amerikanske presidenter (Ballard & Suedfeld

1988: 299-300). Det er derimot lite trolig at sistnevnte faktor er like interessant eller betydningsfull i Norge da landet sjeldent har vært aktiv i krig (Furre 2000: 24; Keilhau 1927)⁷. Det er også sannsynlig at variabler som går på ulike regjeringstyper vil predikere storhetsscoren til norske statsministere, men ikke amerikanske presidenter. Dette belyser viktigheten av faktorene norske historikere vektlegger i sine gjennomganger av statsministere.

Det er trolig en rekke variabler som predikerer en statsministers storhetsgrad, men hva alle disse er og hvordan de skal måles er problematisk å få svar på innenfor rammene av denne analysen. Simonton bruker for eksempel 100 biografiske faktorer for å predikere storhet og 200 for å predikere betydningen av faktorer knyttet til presidenters innsettelse, avsettelse og hva som skjedde mens personen var innsatt. Variablene ble samlet inn over en femårsperiode (1986: 266-268). Begrensningen i denne analysen kan føre til at relevante prediksjonsfaktorer ikke blir identifisert og testet.

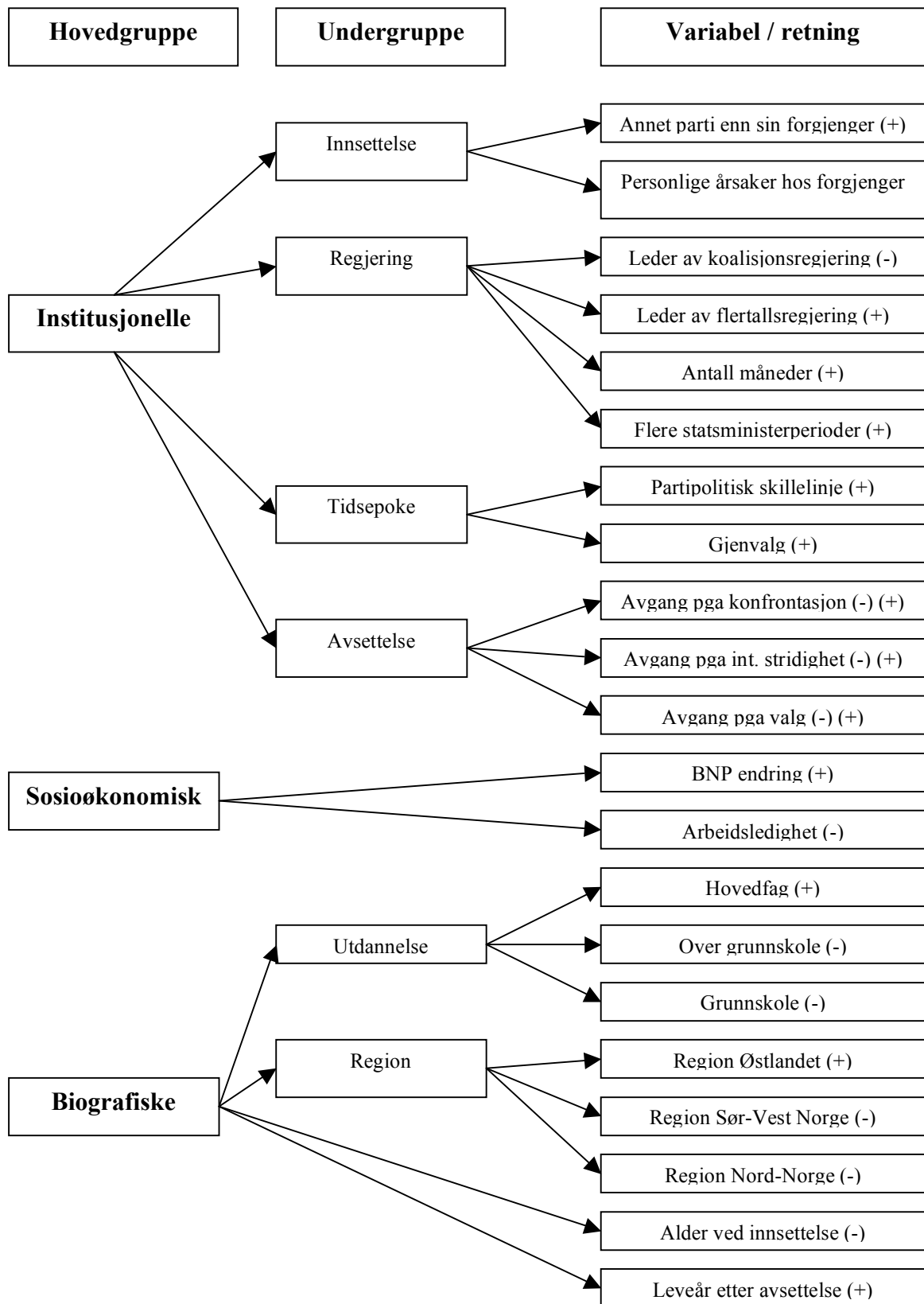
I diskusjonen av uavhengige variabler vil det fokuseres på en bred tilnærming til prediksjon av storhet. Årsaken til dette er det manglende teorigrunnet og studiens eksplorerende utgangspunkt. Fremgangsmåten er i tråd med Simontons studier av amerikanske presidenter. Hans analyser er også utgangspunktet for hovedgruppene av variabler: *institusjonelle, sosioøkonomiske og biografiske* (1986)⁸.

Figur 1 viser hovedgrupper og undergrupper av prediksjonsvariabler denne studien har kommet frem til. Variablenes antatte prediksjon på storhet står her sentralt. I tilfeller der det knyttes usikkerhet til prediksjonen står den antatte prediksjonen først.

⁷ Det kan diskuteres hvorvidt Norge faktisk var passive i 1. Verdenskrig og om landet for eksempel gjennom NATO-medlemskap er aktive i krigføring, men det argumenteres gjerne for at landet kun har vært aktive i Andre verdenskrig. Diskusjonen er likevel utenfor denne analysens hensikt.

⁸ Simonton opererer med to hovedgrupper: *biographical antecedents* og *presidential factors*.

Figur 1 – Variabelgrupper og enkeltvariablers antatte prediksjon på storhetsgrad



3.2.1 *Institusjonelle variabler*

Denne variabelgruppen inneholder faktorer knyttet til institusjonelle forhold som kan predikere graden av storhet blant norske statsministere. Det antas at de institusjonelle variablene vil predikere storhetsgrad bedre enn de sosioøkonomiske og biografiske faktorene. Hovedgruppen er delt opp i fire underkategorier der den første tar for seg variabler relatert til *innsettelse* av statsministeren. Den neste variabelgruppen er knyttet til hva slags *regjering* statsministeren ledet og hvor lenge personen satt ved makten. Den tredje undergruppen er *tidsepoker* mens den fjerde går på årsaker til *avsettelse* av statsministeren.

Undergruppen *innsettelse* består av to variabler der den første er *annet parti enn sin forgjenger*. Simonton bruker denne faktoren uten å utdype antatt prediksjon på storhet. Intuitivt kan det tenkes at det er lettere å ta over fra en statsminister som er fra et annet parti enn sitt eget. Årsaken til dette er at en statsminister fra et nytt parti enklere får gjennomført politiske endringer. Statsministeren kan således bli betraktet som mer handlingskraftig enn hva tilfellet er om personen må videreføre sitt eget partis politikk. På grunnlag av disse antakelsene forventes en positiv prediksjon på storhetsgrad.

Personlige årsaker hos forgjenger. Denne variabelen viser om den nye statsministeren tok over som følge av at forgjengeren trakk seg grunnet personlige årsaker. Det antas at variabelen vil ha negativ prediksjon på grad av storhet fordi slike avganger gjerne kommer uventet og midt i en stortingsperiode. Det kan derfor være vanskelig å komme rett inn og gjøre en god jobb.

Fire variabler ligger til grunn for undergruppen *regjering*. De har hovedsakelig inspirasjon fra norske historikers betraktninger, samt internasjonal litteratur. Variablene har sammenheng med forrige kapittels drøfting av institusjonelle ulikheter mellom land i den forstand at flere av dem ikke er relevante i prediksjon av amerikanske presidenters storhet. Dette gjelder de to første variablene, som viser hvorvidt statsministeren har vært leder for henholdsvis *koalisjons-* og *flertallsregjeringer*. *Antall måneder* omfatter hvor lenge hver statsminister har regjert, uavhengig av forskjellige perioder. *Flere statsministerperioder* betegner hvorvidt en statsminister har vært leder for flere regjeringer, gitt at en eller flere andre statsministere har regjert mellom de ulike periodene.

Leder for koalisjonsregjering. Samarbeid mellom flere partier innebærer at hvert regjeringsparti ikke får gjennomslag for alle sakene sine og at partiet kanskje også må godta en politikk det i utgangspunktet er uenig i. Koalisjonsregjering kan således være et tegn på at statsministeren oppnår lite og blir defensiv i politikken. Handlingskraft, besluttosomhet og styrke er eksempler på faktorer som vektlegges av amerikanske historikere når de ser på en presidents storhetsgrad (Maranell 1970). På den andre siden vil en statsminister som faktisk greier å føre en stabil politikk, og får gjennomslag for saker under slike omstendigheter, kunne bli vurdert positivt i etterkant. Variabelen antas likevel å ha negativ prediksjon på vurderingen av statsministeres prestasjoner.

Leder for flertallsregjering forventes å ha sterk positiv prediksjon på storhet. Denne type regjering gir vanligvis stor grad av handlingskraft og er et tegn på stabilitet. Clark Clifford, som var rådgiver for de amerikanske presidentene Harry Truman og Lyndon Johnson, har presisert at disse to egenskapene er forutsetninger for å bli en stor president: "The public would forgive almost anything in the White House except inaction" (sitert i Murray & Blessing 1994: 41).

Selv om handlingskraft er viktig i både USA og Norge blir det isolert sett mangelfullt å sammenligne landene på dette området da førstnevnte ikke vil ha annet enn flertallsregjeringer. I Storbritannia foretok derimot Theakston og Gill i 2006 en studie med utgangspunkt i Schlesingers metode. Ifølge denne mente respondentene at parlamentarisk flertall var en av de viktigste forutsetningene for å bli vurdert som en stor statsminister i Storbritannia (2006: 207). Koalisjonsregjeringer er ikke vanlig i Storbritannia, men på 1970-tallet måtte for eksempel partiet Labour ha parlamentarisk støtte fra andre partier⁹ (Heidar & Berntzen 1998: 126).

I perioden 1920-35 var det bare mindretallsregjeringer i Norge. Edvard Bull mener datidens regjeringer "[v]ar svake og førte til mye stortingsregjereri (1979: 206). Ifølge boken *Makten og demokratiet* har mindretallsregjeringer vært en tendens som har kommet tilbake etter Arbeidepartiets ettpartistyre ble avsluttet i 1961. Dette har ført til et mer uoversiktlig partilandskap på grunn av det forfatterne betegner som hestehandler mellom partier og partipolitisk samarbeid som varierer fra sak til sak (Østerud, Engelstad & Selle 2003: 84-86).

⁹ Dette tilfellet faller ikke innunder kategorien koalisjonsregjering, men i praksis var det svært likt. Partiet måtte inngå avtaler med andre partier for å få politisk støtte i underhuset.

På tross av den antatt negative prediksjonen av denne type regjeringer på statsministerprestasjoner er for eksempel Gro Harlem Brundtlands statsministerperioder beskrevet som stabile og gode tider (Hansson & Teigene 1992: 323-325). I sine tre perioder var hun aldri leder for en flertallsregjering, og dette kan tyde på at variabelen også kan ha positiv prediksjon hvis regjeringen, på tross av sitt vanskelige utgangspunktet, har vært stabil og hatt gjennomslagskraft. Likevel antas det at Brundtland er et unntak fra normalen og at flertallsregjeringer vil ha positiv prediksjon på storhet.

Antall måneder. Denne variabelen blir regnet for å være blant de viktigste prediksjonsfaktorene på storhetsgrad i USA (Simonton 1986: 273). Den er også sentral i vurderingen av britiske statsministere og forventes å ha sterk positiv prediksjon på bedømmelsen av statsministere i Norge (Theakston & Gill 2006: 205). Betydningen av hvor lenge man sitter ved makten er nært knyttet til hva personen har utrettet, og dermed hva man blir husket for i ettertid (Murray & Blessing 1994: 64 ; Simonton 1986).

En utfordring med *antall måneder* er knyttet til lengre perioder der ett parti har hatt stortingsflertall, slik tilfellet var med Arbeiderpartiet mellom 1945 og 1961. Knut Heidar og Einar Berntsen kaller dette et *ettpartidominert system* (1998: 56). I slike perioder skal det svært mye til for at en flertallsregjering må gå av mellom to valg, og *antall måneder* blir derfor gjerne relativt høyt. Man kan her diskutere hvilket kausalt forhold det er mellom *antall måneder* og storhet, men det er ikke relevant for denne studien.

Flere statsministerperioder. Det at statsministeren får ny tillit av eget parti, og indirekte fra velgerne, er en indikasjon på at arbeidet personen gjorde forrige gang ble verdsatt. Det antas at variabelen vil ha en sterk positiv prediksjon på storhetsscoren.

Undergruppen *tidsepoken* består av to variabler der *partipolitisk skillelinje* er den første. Det antas at denne faktoren vil ha svak negativ prediksjon på statsministervurderingene. Rokkan mener det skjedde et kraftig skille i norsk politikk i 1927, blant annet fordi Venstre ble en del av høyresiden i norsk politikk (1966: 85-89). Dette førte til at "[t]he principal cleavage line was no longer between Left and Right, but between the Socialists and all the offshoots of the parties established before the extension of the suffrage" (Rokkan 1966: 88). Skillene skapte nye konstellasjoner og etter hvert ble flere partier etablert. Dette har trolig medført at

statsministerens rolle har blitt mindre dominerende¹⁰ enn hva tilfellet var da det var færre partier og mer oversiktelige skillelinjer i politikken.

Variabelen *gjenvalg* betegner hvorvidt en statsminister har sittet i posisjon før og etter et stortingsvalg. Variabelen har inspirasjon fra den amerikanske litteraturen og det forventes at den vil ha en sterk positiv prediksjon på statsministerprestasjoner. Ifølge Simonton er *gjenvalg* sentral i prediksjon av storhet selv om den har nær sammenheng med hvor lenge presidenten styrte (1986: 268 og 273).

Variabelen har trolig sterkere prediksjon på amerikanske presidenter enn norske statsministere i den forstand at presidentene oftest blir sittende fire eller åtte år ved makten. Antall år avhenger vanligvis av om presidenten går på et valgnederlag eller trekker seg etter første periode. I Norge er det derimot vanligere å bli innsatt mellom to perioder. Dette er på grunn av de institusjonelle ulikhetene mellom landene, og konsekvensen er at variabelen kan gi mindre styrke på prediksjon av storhet i Norge enn i USA. Det at en statsministers parti får mulighet til å regjere etter et stortingsvalg trenger ikke nødvendigvis å ha prediksjon på personens prestasjon. Et eksempel på dette er Otto A. Blehr som regjerte fra 22. juni 1921 til 6. mars 1923. Det spørres om Blehrs handlinger fra august 1920 til stortingsvalget i oktober samme år førte til at partiet ble gjenvalgt og at han fortsatt var statsminister. Likevel er det i Storbritannia funnet stor sammenheng mellom statsministernes valghistorikk og grad av storhet, noe som styrker argumentet om at variabelen kan predikere storhet i Norge (Theakston & Gill 2006: 205).

Avsettelsesprosessen er den fjerde undergruppen av institusjonelle variabler. Denne variabelkategorien er inspirert fra Simonton (1986) og består av tre variabler. På grunn av usikkerhet knyttet til hvordan faktorene predikerer storhetsgraden til norske statsministere har alle disse variablene alternative hypoteser.

Avgang pga. konfrontasjon måler hvorvidt en statsminister har måttet gå av på grunn av kabinettspørsmål eller mistillitsforslag. Det antas at denne variabelen vil ha negativ prediksjon på storhetsgrad fordi konfrontasjon tyder på manglende tillit, samt at det indikerer

¹⁰ David C. Nice tar også for seg endringer i partisystemer, men han fokuserer mer på hvordan nye partisystemer i USA påvirker en presidents storhet positivt på kort sikt. Etter hvert avgår presidentene som var med på å forme de nye partisystemene forsvinner, og da avtar effekten (1984: 448).

liten grad av handlingskraft. Likevel kan variabelen også vise seg å ha positiv prediksjon. En statsminister som går av på grunn av holdning til en enkelt sak kan tyde på at personen er sterk og ikke viker unna for sine grunnleggende verdier. Det at en statsminister blir avsatt på grunn av mistillit indikerer også at personen er kontroversiell, en egenskap de fleste store amerikanske presidenter ifølge Schlesinger jr. har hatt: “The Greats and Near Greats all took risks in pursuit of their ideals. They all provoked intense controversy. They all, except Washington, divided the nation before reuniting it on a new level of national understanding” (1996: 51).

Variabelen *avgang pga. intern stridighet* i parti/regjering er en indikasjon på om statsministeren hadde manglende tillit. Denne faktoren kan også ha både positiv og negativ prediksjon på storhet. Det at en statsminister var omstridt kan bety at personen ikke gjorde en god jobb, men som nevnt over kan intern stridighet også innebære at statsministeren var kontroversiell og derfor blir vurdert positivt i ettertid.

Avgang pga. valg tyder på at folket var misfornøyd med den jobben statsministeren og regjeringen gjorde, og dermed ikke ga dem fornyet tillit. Det antas at variabelen vil ha negativ prediksjon på avhengig variabel. På den annen side kan det argumenteres for at kontroversielle ledere ikke nødvendigvis var populære da de regjerte, og faktoren kan dermed også ha positiv prediksjon.

3.2.2 Sosioøkonomiske variabler

Denne variabelgruppen inneholder to variabler som omhandler nasjonale mål på den makroøkonomiske utviklingen i landet.

BNP endring er en faktor norske historikere vektlegger i sine omtaler av historiske perioder og statsministere (se for eksempel Hansson & Teigene 1992: 323-324). Simonton påpeker at den økonomiske utviklingen påvirker respondentenes evalueringer av amerikanske presidenter (1986: 279). Økning i BNP antas å ha en sterk positiv prediksjon på respondentenes vurdering av storhetsgrad. I boken *Makt og Mannefall – historien om Gro Harlem Brundtland* poengteres det at hun som statsminister var uheldig med de økonomiske konjunktorene, og at det derfor er en prestasjon at hun likevel greide å styre landet på en god måte (Hansson & Teigene 1992: 324). Dette indikerer at variabelen blir vektlagt av respondentene og at den predikerer statsministerprestasjoner.

Arbeidsledighet antas å ha sterk negativ prediksjon på storhetsgrad. Hansson og Teigene skriver om Einar Gerhardsen at han "[f]remfor alt [var] en statsminister som aldri styrte landet ut i arbeidsløshet" (1992: 323). Her kobles statsministerens prestasjon direkte opp mot arbeidsløshet. På tross av at det er en vanlig antakelse at flere nasjonale og internasjonale forhold påvirker antall sysselsatte er beskrivelsen et tegn på at variabelen er sentral for respondentenes vurderinger av norske statsministere, og at den kan ha prediksjonskraft i modellen.

3.2.3 Biografiske variabler

Denne variabelgruppen består av åtte prediksjonsfaktorer der *utdannelse* og *geografi* er to undergrupper. Ifølge den internasjonale litteraturen har biografiske faktorer lite å si for graden av storhet blant amerikanske presidenter (Murray & Blessing 1994: 23; Schlesinger jr. 1997; Simonton 1986: 275 og 277; Theakston & Gill 2006). Likevel bør variablene testes i norsk kontekst.

Undergruppen *utdannelse* består av variablene *grunnskole*, *over grunnskole* og *hovedfag*. De to førstnevnte antas å ha svak og negativ prediksjon mens det antas positiv prediksjon av *hovedfag* på storhetsgrad. Det er trolig at en statsminister med hovedfag vil ha et bedre utgangspunkt for å bli vurdert som stor enn en regjeringsleder med liten eller ingen utdannelse. Likevel kan det tenkes at enkelte statsministere lettere kan bli populære blant sosiale grupper som kan identifisere seg med regjeringslederen. For eksempel blir Einar Gerhardsen i boken *Einar Gerhardsen – en politisk biografi* beskrevet som arbeidergutt og en mann av sin tid (Olstad 1999: 427). Dette kan knyttes opp mot ikonet som landsfader og medføre at Gerhardsen får en "Askeladden-effekt" som gjør at han blir vurdert høyere for å være en folkets mann.

Undergruppen *geografi* ser på prediksjonseffekten av hvilken landsdel statsministeren har representert på Stortinget, men det er usikkerhet knyttet til variabelens prediksjonen på storhet. Stein Rokkan går i boken *Stat, nasjon, klasse* inn på ulike typer konflikter og maktkamper mellom sentrum og periferi. Han argumenterer for at utviklingen i Norge går mot en sterkere konsentrasjon av elitearenaer i Oslo selv om det også blir klarere distinkte maktressurser utenfor sentrum (1987: 232-234). I denne sammenheng kan det argumenteres for at det er vanskeligere å bli positivt vurdert jo lenger borte fra sentrum statsministeren har sin geografiske tilhørighet. Problemene med å påvirke det etablerte sentrum vil i så måte føre

til at både Nord-Norge og Sør-Vestlandet ha negativ prediksjon, og førstnevnte i større grad enn sistnevnte. *Østlandet* vil derimot ha positiv prediksjon på storhet.

Variabelen *alder ved innsettelse* antas å ha svak negativ prediksjon på storhetsgrad. Muligheten til å regjere over lengre tid og flere perioder øker isolert sett jo yngre man er ved første innsettelse. Denne variabelen er også brukt av Theakston og Gill i studien av britiske statsministere. *Alder ved innsettelse* hadde her svært lav korrelasjon med storhetsgrad (2006: 202-204).

Kelly og Lønnstrøm hevder at presidenter, etter de har forlatt stillingen, kan påvirke historikere og samfunnsviteres vurdering av dem. Dette kan foregå gjennom for eksempel selvbiografier og aktiv deltakelse i media (2003: 632-633). Flere *leveår etter avsettelse* vil øke mulighetene for statsministerne til å påvirke sin egen rangering, og det antas at variabelen vil ha positiv prediksjon på storhetsgrad. Hansson og Teigene bruker Einar Gerhardsen som eksempel på statsministeres mulighet til å påvirke sitt ettermæle:

”Historiens bilde av ham - som han selv i høyeste grad bidro til å skape gjennom sine mange bøker på slutten av sitt liv - er et bilde av mannen som med myk hånd sto i spissen for flokken som løftet et helt folk fram fra fattigdom og velstand” (1992: 323).

3.3 Oppsummering av hypotesekapittel

Dette kapitlet består av en variabeldiskusjon der avhengig variabel, samt undergrupper og variabler innenfor hovedgruppene *institusjonelle*, *sosioøkonomiske* og *biografiske* variabler ble gjennomgått. Inspirasjonen for variablene, og presentasjon av hypoteser om variablenes antatte prediksjon på grad av storhet, sto her sentralt. Gruppen av institusjonelle variabler ble ansett som mest relevante mens den biografiske gruppen ble vurdert som mindre sentral. Oppgaven går nå over til å diskutere dataene som ligger til grunn for studien, samt operasjonaliseringen av dem. Metodiske problemstillinger relevante for den kommende analysen vil også bli gjennomgått.

4. DATA, OPERASJONALISERING OG METODE

Dette kapittelet består av tre hoveddeler. Den første tar for seg datagrunnlaget for Cordtsen-undersøkelsen (for utfyllende datasett se appendiks A). Dette er en viktig gjennomgang fordi verdiene på avhengig variabel defineres med utgangspunkt i denne undersøkelsen. Det blir her også diskutert mulige problemer knyttet til respondentenes objektivitet. Den andre hoveddelen presenterer og drøfter operasjonaliseringen av variablene som ligger til grunn for studien. Metodiske temaer i forhold til den kommende analysen blir diskutert i den tredje hoveddelen. Formålet med denne studien er å finne ut hvilke faktorer som predikerer graden av storhet blant norske statsministere i perioden 1905-1997.

4.1 Cordtsen-studien

I september 2007 oversendte Kjell Cordtsen samtlige originaldokumenter som var lagt til grunn for Verdens Gang-artikkelen i 1997, deriblant fremgangsmåten for datainnsamlingen samt respondentenes svar og utdypende vurderinger. Originaldokumentene ble så kopiert, og det ble bygget opp en ny datafil med verdier for avhengig variabel. Det ble ikke observert mangler i gjennomgangen, og det var samsvar mellom Cordtsens tall og denne studiens utregninger.

I en e-post 22.mars 2008 utdypet Cordtsen valgene han foretok i forbindelse med undersøkelsen. Når det gjaldt utvelgelsen av respondenter hadde journalisten selv studert historie og statsvitenskap, og følte derfor han hadde relativt god oversikt over fagmiljøene. For å være sikrere på sammensetningen av svarpanelet søkte han også råd hos flere personer innenfor disse akademiske miljøene. Etter å ha satt opp et representativt panel ble de aktuelle personene kontaktet per telefon og presentert for Schlesinger jr.-undersøkelsen i New York Times Magazine fra 1997. De ble videre bedt om å delta i en tilsvarende norsk undersøkelse. Deretter fikk respondentene tilsendt et brev med listen over Norges statsministere etter 1905 der datoer for tiltredelse og avgang for hver av dem var presisert. Cordtsen la også ved en oversikt over alle personene som var blitt forespurt deltakelse. Han håpet dette ville gi høyere svarprosent fordi respondentene ved å se listen kunne føle seg i godt selskap. 20 av 33, eller 60 prosent, av de spurte valgte å delta i undersøkelsen (for fullstendig oversikt over respondentene se appendiks B). Ifølge Cortsen var dette mer enn forventet. Respondentenes hovedargumenter mot å delta i studien var for det første at det var vanskelig å gi statsministerne score uten å gå dypere inn på bakgrunnen for disse. Dernest mente flere at

det var problematisk å foreta en vurdering av statsministerne når ulike kriterier ville bli lagt til grunn. For det tredje ga noen respondenter uttrykk for at de hadde mangelfull tid til å sette seg skikkelig inn i materialet. Argumentene mot deltakelse samsvarer med canadiske respondents reaksjoner på Ballard og Suedfelds lignende studie i Canada (1988: 295).

Åtte av til sammen 25 statsministere har regjert maksimum ett år, og på grunnlag av det korte tidsperspektivet bedømte ikke alle respondentene samtlige statsministere. Likevel går antall statsministervurderinger aldri under 16, og dette innebærer at hver statsminister er vurdert av minimum 80 prosent av respondentene. Derfor blir alle regjeringslederne inkludert i analysen. I Ballard og Suedfelds undersøkelse (Ballard & Suedfeld 1988) ville svært få respondenter vurdere statsministerne som satt over kortere perioder, og derfor ble disse tatt ut av deres analyse.

Cordtsen-studien inneholdt seks kategorier: stor (seks), betydelig, over middels, middels, under middels og dårlig (en), noe som avviker fra Schlesingers fem kategorier¹¹. Dette valget ble gjort fordi Cordtsen antok at respondentene ville plassere mange statsministere i middelskategorien, og derfor ønsket han en større nyansering i rangeringene.

¹¹ Schlesinger valgte å ha verdiene 1-5 der 5=great, 4=near great, 3=average 2=below average og 1=poor. Dette er også verdiene Schlesinger jr. opererer med.

Tabell 1 viser statsministere, perioden de var innsatt og gjennomsnittsscoren de fikk av respondentene.

Tabell 1 – Statsministerne og deres storhetsscore

Statsminister	Statsministerperioder	Score
Christian Michelsen	1905-1907	5.75
Jørgen Løvland	1907-1908	2.63
Gunnar Knudsen	1908-1910, 1913-1920	5.25
Wollert Konow	1910-1912	1.84
Jens Bratlie	1912-1913	2.18
Otto B. Halvorsen	1920-1921, 1923	2.78
Otto Blehr	1921-1923	2.58
Abraham Berge	1923-1924	2.17
Johan L. Mowinckel	1924-1926, 1928-1931, 1933-1935	3.85
Ivar Lykke	1926-1928	2.21
Christopher Hornsrud	1928	3.56
Peder Kolstad	1931-1932	1.80
Jens Hundseid	1932-1933	1.45
Johan Nygaardsvold	1935-1940, 1940-1945	4.50
Einar Gerhardsen	1945, 1945-1951, 1955-1963, 1963-1965	5.85
Oscar Torp	1952-1954	3.00
John Lyng	1963	4.18
Per Borten	1966-1970	3.50
Trygve Bratteli	1971-1972, 1974-1975	3.85
Lars Korvald	1973	3.16
Odvar Nordli	1976-1980	2.90
Gro H. Brundtland	1981, 1986-1989, 1991-1996	4.95
Kåre Willoch	1981-1983, 1983-1986	4.40
Jan P. Syse	1990	2.21
Thorbjørn Jagland	1997	1.68
Snitt		3.30

Av tendenser i tabellen kan det trekkes frem at gjennomsnittsscoren til samtlige statsministere er 3.3. Frem til Gerhardsens første regjering i 1945 er snittet 3.0 mens det er 3.6 for etterkrigsperioden. Ser man på ulike perioder går det frem at foruten Gunnar Knudsen (5.25) er det kun to av ti statsministere, Johan L. Mowinckel (3.85) og Christopher Hornsrud (3.56), som oppnår et bedre gjennomsnitt enn 3.0 mellom årene 1910 og 1935. Dette er påfallende få sammenlignet med hele den observerte perioden. Mellomkrigstidens problemer med ustabile og kortvarige regjeringer, samt økonomisk nedgangstid, kan være årsaker til tendensen. Dette vil bli diskutert mer inngående i analysekapittelet.

I utformingen av analysens datasett ble det gjort et valg mellom to alternative måter å dele inn enhetene på. Enten kunne hver statsminister være utgangspunkt for kun én observasjon i et minimalistisk datasett, eller så kunne ulike regjeringer være utgangspunkt for hver

observasjon i et maksimalistisk datasett. I alternativ nummer to ville enhver ny kombinasjon av a) statsminister, b) parti i regjering og c) parlamentarisk grunnlag bli brukt som utgangspunkt for observasjoner. Parlamentarisk grunnlag innebærer her overganger mellom flertalls- og mindretallsregjeringer og ettparti- og koalisjonsregjeringer. Mens det minimalistiske datasettet ser på hver statsminister under ett, ville det maksimalistiske i større grad predikert regjeringsperioder. Sistnevnte alternativ kunne for eksempel vist at Einar Gerhardsens første periode ble bedre predikert enn hans andre periode.

Denne analysen går videre med det minimalistiske alternativet, altså kun én observasjon per statsminister. Respondentenes svar, som er utgangspunkt for verdiene på avhengig variabel, er gitt med utgangspunkt i de ulike statsministeres samlede prestasjoner, ikke for hver statsministers ulike regjeringer og perioder. Følgelig kan det argumenteres for at det ville blitt upresist å benytte seg av observasjonsinndelingen i det maksimalistiske datasettet som utgangspunkt for analysen. Metoden samsvarer med den Theakston og Gill brukte i studien av britiske statsministere (2006: 197). Det er også slik de fleste studiene av presidentprestasjoner i USA blir gjennomført.

4.2 Objektive respondenter?

Respondentenes vurderingen av norske statsministere, som er grunnlaget for verdiene på storhetsscore, kan påvirke samtidens og fremtidens oppfatninger av landets statsministere. Dette er interessant fordi respondentene selv bestemte kriteriene for storhet. Det kan derfor argumenteres for at respondentenes vurderinger forteller oss mer om hvordan disse personene knytter politiske egenskaper til statsministere enn de predikerer faktiske statsministerprestasjoner. Storhetsscoren avhenger således i større grad av personene som observerer enn statsministerne som blir observert (Simonton 1986: 260). Dette argumentet støttes også av Theakston og Gill som mener at "it reveals more about the views and values of the academic respondents than about actual presidential performance" (2006: 194). Det vil her bli diskutert fire utfordringer knyttet til respondentenes objektivitet i denne type studier.

For det første kan respondentenes politiske ståsted påvirke deres vurderinger. Generelt er akademikere mer venstreorienterte enn den øvrige befolkningen, og dette kan ha innvirkninger på bedømmelsene av statsministerne (Murray & Blessing 1983: 545). Ifølge

James Taranto¹² reflekterer venstredreiningen blant akademiske respondenter i USA denne type undersøkelser (2005: 1). Valgundersøkelsen 2005 er en indikasjon på at norske akademikere også er generelt venstreorienterte. Den viser at 30 prosent av personer med høyere utdanning stemmer på Arbeiderpartiet mens 15 prosent stemmer på Sosialistisk Venstreparti¹³ (Aardal, Høstmark, Lagerstrøm & Stavn 2007: 13). Som den 37. amerikanske presidenten Richard Nixon uttalte på *Meet The Press*¹⁴ 10. april 1980: “History will treat me fairly. Historians probably won’t. They are mostly on the left” (siteret i Schlesinger jr. 1997: 180). De britiske respondentene i Theakston og Gills undersøkelse viste seg å bekrefte antakelsen (2006: 199 og 208). Jeffrey E. Cohen er derimot uenig i dette og referer til analyser av respondenter som viser at vurderingene ikke blir påvirket av deres politiske ståsted (2004: 2 og 3).

En andre utfordring knytter seg til respondentenes integritet. Arthur M. Schlesinger jr. foreslo for eksempel hvilke grep Bill Clinton burde foreta for å få en høyere status blant historikere i forbindelse med presidentvalget i 1996 (Bose & Landis 2003: 73). Selv om Schlesinger selv ikke selv har deltatt i studier av presidenters storhet, kan rollene som rådgiver og uavhengig historiker være problematisk. I denne sammenheng kan man heller ikke se bort fra at respondentene har egne politiske agendaer i forbindelse med vurderingene. Dette kan være for å skape blest om egen person eller at respondentene av andre grunner ikke er objektive i bedømmelsen av statsministerne. Slike forhold kan føre til at det stilles spørsmål omkring verdien av denne type studier. Bose og Landis ser likevel på den mulige påvirkningen respondentene har på resultatene som et positivt og interessant element i forskningen og den tilhørende diskusjonen rundt studier av presidentprestasjoner (2003: 41).

For det tredje er alle respondentene i Cordtsen-undersøkelsen menn¹⁵. Dette kan påvirke hvilke kriterier de ser etter når statsministerne vurderes. Problemet er diskutert i den internasjonale litteraturen, og Bose og Landis utdyper konsekvensen av dette: ”Male scholars evaluating male presidents. Hence they are tested by so-called male qualities” (2003: 93). På tross av et lite utvalg viste deg seg at kvinnelige respondenter i Storbritannia generelt

¹² James Taranto er redaktør for Opinionjournal.com og medredaktør i *Presidential Leadership: Rating the Best and the Worst in the White House*.

¹³ Hva som er høyre- og venstresiden i norsk politikk kan diskuteres, men uansett er flertallet av de med høyere utdanning politisk venstreorientert.

¹⁴ TV-program på NBC som ofte har verdensledere og amerikanske toppolitikere som sine gjester (<http://www.msnbc.msn.com/id/3403008/>).

¹⁵ Tre kvinnelige akademikere ble spurt om å delta i undersøkelsen, men takket nei.

vurderte statsministerne mer positivt enn mannlige (Theakston & Gill 2006: 201). Likevel er det gjort funn fra studier i USA som tyder på at kjønn i liten grad påvirker respondentenes vurderinger (Murray & Blessing sitert i Simonton 1986: 262).

Et fjerde utfordrende aspekt er at respondentene kan bli påvirket av hvilke statsministere som tradisjonelt sett har blitt mye omtalt i positiv eller negativ forstand, både i media og av historikere. Disse statsministerne er det gjerne foretatt mer omfattende studier av enn dem som ikke markerte seg i så stor grad. Som en kritiker av amerikanske presidentrangeringer har uttalt: "Certain leaders are great because we call them great" (sitert i Crockett 2002: 241). I denne type studie der det skal foretas retrospektive vurderinger av statsministere, og da spesielt de langt tilbake i tid, vil det også være fare for at respondentenes vurderingsgrunnlaget varierer og at validiteten i resultatene svekkes (Steenbergen & Marks 2007: 349).

På tross av disse fire mulige utfordringene er det gjort funn som viser at respondentenes kjønn, alder, geografisk bosted og profesjonsstatus i liten grad påvirker vurderinger av presidenters storhet (Murray & Blessing sitert i Simonton 1986: 262). Det er i tillegg stor grad av samsvar mellom presidentstudier som har inkludert forskjellige respondenter. Korrelasjonen mellom respondenters vurderinger av presidentprestasjoner i ulike studier er vanligvis på over 0.9. Dette gjelder også for dem som opererer med forskjellige skalaer for å måle storhet (Simonton 1986: 262). Funnene styrker verdien av denne type studier fordi det øker sannsynligheten for at like kriterier legges til grunn for vurderingene. Det antas dessuten at historikerne og samfunnsviterne i denne studien har gode kunnskaper om norsk historie, og at de ved å være med på en statsministervurdering faktisk gjennomfører oppgaven på en mest mulig objektiv måte. De nevnte forholdene styrker bruken av respondentenes svar som utgangspunkt for å analysere hvilke faktorer som predikerer graden av storhet blant norske statsministere.

I tillegg vil residualene si noe om hvor godt de ulike statsministerne blir predikert av analysens endelige modell. Man vil da kunne fastslå hvor stor innflytelse eksterne variabler har på de enkelte statsministernes storhetsscore. Residualdiskusjonen vil også gå inn på alternative prediksjonsvariabler til enhetene som blir dårligst predikert av modellen. Således vil ikke statistisk prediksjon av statsministerprestasjoner tolkes isolert.

I denne type studier må respondenter velges ut fra hvem som har best grunnlag for å vurdere statsministerprestasjoner. Ballard og Suedfeld mener mål av storhet på grunnlag av akademikerens bedømmelser er viktig av to grunner (1988: 300-301). For det første må enhver evaluering av lederprestasjoner til en viss grad være et spørsmål om subjektive tolkninger og oppfatninger på grunn av kompleksiteten i slike studier. For det andre er eksperters vurderinger mer valide mål på statsministeres prestasjoner enn det alternativene er. For eksempel er det trolig at denne gruppen har mer inngående kjennskap til Norges statsministere enn hva tilfellet er i den allmenne befolkningen. Hvorvidt akademikerne er mer objektive i sine vurderinger enn den øvrige befolkningen er likevel uvisst.

4.3 Operasjonalisering av variabler

Som vist i hypotesekapittelet blir det i utgangspunktet brukt 21 variabler til å predikere de 25 statsministernes prestasjoner (for utfyllende datasett og koding se appendiks 1). Årsaken til valget om å inkludere et så høyt antall variabler har sammenheng med målet om en bred tilnærming til problemstillingen. Tatt i betraktning at dette er første gang en slik analyse gjennomføres i Norge har man få forutsetninger for å vite hvilke variabler som vil predikere statsministerprestasjoner. Det kan dessuten ikke forventes at alle variablene vil ha signifikant prediksjon på storhet. Senere i kapittelet blir det etablert kriterier for hvilke variabler som blir valgt ut til den endelige modellen. Hvis det mot formodning viser seg at antall variabler er et problem vil dette bli diskutert i analysekapittelet.

Tre hovedgrupper av uavhengige variabler ligger til grunn for analysen: *institusjonelle*, *sosioøkonomiske* og *biografiske*. Alle dummyvariablene i gjennomgangen har verdi 1 som bekreftende.

4.3.1 Institusjonelle faktorer

Data som danner grunnlag for de institusjonelle variablene, bortsett fra *partipolitisk skillelinje*, er hentet fra Samfunnsveven¹⁶. Dette er en informasjonstjeneste utviklet av Norsk Samfunnsvitenskapelig Datatjeneste AS (NSD).

Annet parti enn sin forgjenger er en dummyvariabel som er operasjonalisert ved å se på hvilke statsministere som en eller flere ganger har tatt over stillingen etter en statsminister

¹⁶ www.samfunnsveven.no.

fra et annet parti enn sitt eget. Gunnar Knudsen, Einar Gerhardsen og Gro Harlem Brundtland har alle overtatt stillingen en eller to ganger fra en statsminister fra eget parti, og en eller to ganger fra en person fra et annet parti. De aktuelle enhetene blir kodet på grunnlag av den perioden de satt lengst som statsminister da det er trolig at denne i størst grad påvirker respondentenes helhetlige vurdering. For eksempel var Gunnar Knudsen statsminister i over seks år etter å ha overtatt fra Høyres Jens Bratlie i 1913, men regjerte kum i underkant av to år etter å ha overtatt fra sin partikollega Jørgen Løvland i 1908.

Personlige årsaker hos forgjenger er en dummyvariabel som er operasjonalisert ved å se om den aktuelle statsministeren ble innsatt som følge av død eller ønske om avløsning hos forgjengeren.

Operasjonaliseringen av variablene leder for koalisjonsregjering og leder for flertallsregjering er foretatt ved å se på hvilken type regjering statsministerne har styrt. Gunnar Knudsen, Einar Gerhardsen og Kåre Willoch har styrt med ulike kombinasjoner av ettparti- og koalisjonsregjeringer, samt flertalls- og mindretallsregjeringer. I disse tilfellene kodes statsministerne etter hvilken regjeringstype de ledet over lengst periode.

Variabelen *antall måneder* er operasjonalisert litt annerledes enn hva tilfellet er i mange lignende studier i USA. Det er vanligvis kun spesielle omstendigheter som gjør at amerikanske presidenter ikke fullfører sin periode, og *antall år* brukes gjerne i disse studiene¹⁷. I Norge er det derimot betydelige forskjeller i regjeringstiden. Derfor operasjonaliseres variabelen gjennom *antall måneder som statsminister* målt fra den måneden statsministeren ble innsatt til og med den måneden personen ble avsatt. Den ulike operasjonaliseringen har sammenheng med de institusjonelle ulikhetene diskutert i kapittel 2. Regjeringstiden varierer fra Christopher Hornsrud som satt en måned, til Einar Gerhardsen som var statsminister i totalt 209 måneder.

Partipolitisk skillelinje er operasjonalisert ved at det skilles mellom statsministerne til og med 1927 (verdi 1), og dem som har blitt innsatt etter dette året. Kodingsvalget er foretatt i henhold til Stein Rokkans modell om partipolitisk skille i norsk politikk (1966).

¹⁷ Dette varierer likevel. Simonton (2001:3) deler for eksempel ett år opp i tideler for å få et mer korrekt mål på innsettelsesperioden til amerikanske presidenter.

Gjenvalg er operasjonalisert ved å se på om samme person har sittet som statsminister før og etter et stortingsvalg. Knudsen, Lykke, Gerhardsen og Brundtland har både blitt gjenvalgt og ”tapt” valg. Foruten Lykke satt likevel alle personene lengst som statsministere i periodene før og etter gjenvalget. Det er sannsynlig at deres prestasjoner i denne lengste perioden har hatt større betydning enn i perioden da de tapte et stortingsvalg.

Når det gjelder årsaker til *avsettelse* så har Gunnar Knudsen, Otto B. Halvorsen, Einar Gerhardsen, Trygve Bratteli, Gro Harlem Brundtland og Kåre Willoch gått av på grunn av ulike årsaker. I disse tilfellene vil avgangsårsaken etter den lengste innsettelsesperioden bli brukt. Det er sannsynlig at det er prestasjonene og hendelsene i denne perioden som i størst grad ligger til grunn for respondentenes vurderinger¹⁸. De aktuelle variablene er *avgang pga. henholdsvis konfrontasjon, intern stridighet og valg*.

4.3.2 Sosioøkonomiske faktorer

Det lange tidsperspektivet i datasettet gjorde det utfordrende å finne tilgjengelige data på de sosioøkonomiske variablene, samt å kode disse.

I utgangspunktet var det tenkt at analysen skulle inneholde en variabel med gjennomsnittlig årlig endring i BNP per innbygger, regnet i prosent, i hver statsministers samlede regjeringsperiode. Beslutningen om at variabelen ikke kunne ta hensyn til innbyggere har sammenheng med at det blir matematisk komplisert å foreta utregninger for BNP per innbygger for statsministere som har regjert i flere ulike perioder. Grunnen til dette er at det i tallmaterialet fra Statistisk sentralbyrå (SSB) ikke er regnet ut BNP per innbygger fra årene før 1970¹⁹. Dermed måtte tallene fra perioden 1905-1970 regnes ut manuelt på grunnlag av data på middelfolkemengden²⁰ for å kontrollere for innbyggere. Dette viste seg å være komplisert og upresist.

Valget falt på variabelen årlig prosentvis endring i BNP i hver statsministers samlede regjeringsperiode. Det er viktig å få endringstallenes prediksjon på statsministeres storhet da BNP, med unntak fra tiden omkring Andre verdenskrig, stort sett har økt i hele perioden 1905-1997. I forbindelse med Andre verdenskrig foreligger det ingen BNP-tall for årene

¹⁸ Denne antakelsen ble bekreftet da respondentenes vurderinger ble gjennomgått i Cordtsens originaldokumenter.

¹⁹ Dette ble bekreftet i en e-post 7. desember, og SSB sendte da data på *middelfolkemengde* i perioden 1905-1970.

²⁰ For hvert femte år har SSB nye tall for middelfolkemengden, og en statsministers periode krysser ofte av disse. Dette gjør utregningene kompliserte.

1939-1945, og data beregnes på grunnlag av gjennomsnitt av verdiene for femårsperioden før og etter henholdsvis 1940 og 1945.

Gjennomsnittlig *arbeidsledighet* måles med utgangspunkt i årene en statsminister har regjert lengre enn halve kalenderåret. For eksempel blir arbeidsledigheten i Otto B. Halvorsens første regjering, som varte fra 21.juni 1920 til 22.juni 1921, kun målt gjennom året 1920. For statsministere som satt mindre enn seks måneder til sammen vil de bli målt gjennom arbeidsledigheten for det aktuelle året, eller eventuelt det året de satt lengst.

Når det gjelder *arbeidsledighet* så har utfordringen vært å finne tallmateriale, samt at de ulike dataene skal være sammenlignbare for hele perioden 1905-1997. SSBs *Historisk Statistikk 1978* og *1994* har vært kilder for tall på arbeidssøkende personer i perioden etter 1960. Disse tallene er utarbeidet i samarbeid med ulike trygdeetater i perioden. For perioden før 1960 var tallmaterialet derimot svært mangelfullt²¹. Ola H. Grytten ved NHH har forsket mye på arbeidsmarkedsutvikling, og det er hans tall²² som benyttes for årene 1905-1959. ”As far as possible I have applied the same definitions of unemployment as the Central Bureau of Statistics has applied for the post-World War II period” (Grytten 1994: 2).

4.3.3 Biografiske faktorer

Data til de biografiske variablene er hentet fra NSDs ”Politikerarkivene 1814-2001”²³ og regjeringens nettsider²⁴.

Undergruppen *utdannelse* er operasjonalisert ved bruk av kategoriene *grunnskole*, *over grunnskole* og *hovedfag*. Den første kategorien inneholder statsministere som enten har avbrutt grunnskolen eller ikke har utdannet seg mer enn dette. *Utdanning over grunnskole* innebærer all form for utdanning som overgår grunnskole, men som ikke er fullført hovedfag eller tilsvarende. Utdanning tilsvarende *hovedfag* innbefatter statsministere som har tatt hovedfag eller mer enn dette.

²¹ SSB informerte i en e-post 6.november at de kun hadde oversikt over antall registrerte arbeidsledige i perioden før 1972.

²² Noen av disse dataene er relativt nye, og varierer litt fra dem som er hentet fra Gryttens bok. Grytten sendte 23.11.07 en e-post med oppdaterte tall.

²³ For full oversikt se:

<http://extweb3.nsd.uib.no/data/polsys/Index.cfm?ArkivNr=12&Institusjonsnummer=1&TypeTekst=98&Uttaksnummer=18&Sprak=>

²⁴ www.regjeringen.no

I utgangspunktet er dette en variabel på ordinalnivå, men den er her gjort om til to dummyvariabler der *hovedfag* er referansekategori. På tross av endringer i utdanningssystemet²⁵ fra midten av 1800-tallet har lengden på grunnskolen generelt sett endret seg i sammenheng med utviklingen i samfunnet. Når det gjelder høyere utdanning så har fire av de fem første norske statsministre og seks statsministere etter hverandre i perioden 1928-1951 ikke har slik utdanning. Dette tyder på at det ikke er stor økning i andelen av høyere utdannede i denne perioden. Derimot er det en svakhet i dataene at utdanningsnivået etter 1960 har utviklet seg kraftig²⁶. De fire siste statsministre i datasettet har alle minst hovedfag, og det er også sannsynlig at de fleste av de kommende statsministre vil ha dette.

Undergruppen *geografi* innebærer de tre variablene *representant fra Sør-Vestlandet, Nord-Norge og Østlandet* der sistnevnte fungerer som referansekategori. John Lyng har vært representant for både Østlandet og Nord-Norge²⁷, men plasseres innunder Nord-Norge fordi han var representant for denne landsdelen over lengst periode.

Operasjonaliseringen av *alder ved innsettelse* har ingen problematiske aspekter ved seg.

Leveår etter avsettelse fører til utfordringer knyttet til statsministere som fortsatt lever. For dem²⁸ er verdiene satt ut fra gjennomsnittlig levealder i den norske befolkning²⁹. Ifølge SSB var denne i 2006 79,1 år for menn og 82,7 år for kvinner. Hvis den tidligere statsministeren allerede har oversteget dette nivået blir verdiene satt med utgangspunkt i at personen har levd frem til februar 2008.

Av operasjonaliseringen ser man at noen variabler er svakere enn andre. Enkelte statsministere har for eksempel gått av på grunn av ulike årsaker, og dette gjør det utfordrende å kode dem. I tillegg er det en svakhet i *utdanning* at det generelle utdanningsnivået er blitt høyere i Norge, spesielt siden 1960-tallet. Validitet vil bli nærmere diskutert senere i dette kapittelet.

²⁵ Den femårige folkeskolen, som er forgjengeren til grunnskolen, ble innført i 1889, men ble endret i 1920 til å være syvårig. I 1969 ble navnet endret til grunnskole, og skolen ble da niårig (Brøgger 1983: 95 og 128). Det har også vært endringer i den tidligere framhaldsskolen og realskolen.

²⁶ Se for eksempel <http://www.ssb.no/histstat/aarbok/hf-040240-190.html>

²⁷ Nord-Norge innebærer alt fra Sør-Trøndelag og nordover. Bortsett fra Otto A. Blehr, som både var representant fra Sør-Trøndelag og Nordland, har samtlige av de andre innenfor denne kategorien representert Sør- eller Nord-Trøndelag.

²⁸ Odvar Nordli, Gro Harlem Brundtland, Kåre Willoch og Thorbjørn Jagland.

²⁹ (<http://www.ssb.no/emner/02/02/10/dode/>).

4.4 Kvantitativ metode

Komparativ samfunnsvitenskap er vanligvis delt i to forskningsstrategier: hypotesetesting gjennom den kvalitative case-orienterte tilnærmingen og den kvantitative variabel-orienterte (Ragin 1987: 16-18). Førstnevnte har som hovedmål å skape kunnskap innenfor historisk-kontekstuelle rammer og er passende når man foretar inngående studier av få case. (Grimen 2000: 204-205 og; Johannessen & Tufte 2002; Ragin & Zaret 1983).

I denne studien er formålet å teste hvilke faktorer som predikerer grad av storhet blant samtlige norske statsministere i perioden 1905-1997. I tillegg skal det fokuseres på hvor godt de ulike statsministerne blir predikert av variablene som er inkludert i modellen. De nevnte faktorene er i samsvar med den kvantitative forskningsstrategien.

4.5 Multivariat regresjon

For å studere hvilke faktorer som predikerer graden av storhet blant norske statsministere vil minste kvadratsums metode (OLS) benyttes. I motsetning til korrelasjonsanalyse, som kun viser samvariasjon mellom variabler, åpner dette analytiske redskapet opp for å påvise retningen og styrken av en prediksjon. I denne analysen brukes tre hovedgrupper av variabler. Den multivariate regresjonsanalysen gjør det da mulig å "[i]dentifisere den samlede påvirkningen av gruppene som helhet så vel som den unike effekten av enkeltvariablene innenfor hver av gruppene" (Midtbø 2007: 97). I statistisk forstand gjør dette at det blir lettere å skille mellom gode og dårlige prediksjonsvariabler.

Analysen består av én avhengig metrisk variabel, storhetsgrad, som kan predikeres av flere uavhengige metriske variabler. Ifølge Hair et.al. (1998: 14) er multivariat regresjon passende under slike forhold.

4.5.1 Mål og tester i regresjon

Når man skal analysere funn fra regresjonsanalyser, er det nødvendig å forklare mål og tester som vil bli anvendt. De som er relevante for enkeltvariablene i analysen vil bli gjennomgått først. Deretter rettes fokus mot mål og tester som omhandler selve modellene.

Signifikanstester brukes for å kartlegge hvilke prediksjonseffekter som er hevet over usikkerhet. Det finnes ulike grenser for signifikansnivå, og i denne analysen blir det kritiske nivået satt til ti prosent. Dette innebærer at variabler som har p-verdier under 0.10 fra de

gruppevise regresjonsanalysene vil bli analysert videre. P-verdien angir det nøyaktige signifikansnivået og sannsynligheten for at nullhypotesen, som sier at det ikke er noen effekt, blir forkastet når den faktisk er sann. Årsaken til at signifikansnivået i denne analysen ikke settes til fem prosent, som er mer vanlig, er at variablene er lite robuste i den forstand at de ikke bygger direkte på tidligere studier. Man vet derfor lite om deres prediksjonseffekt, noe som har sammenheng med at det er første gang en slik analyse gjøres i Norge. En slik utforskende studie gjør at det kan vises litt mer fleksibilitet til signifikansnivået. Konsekvensen av å sette et høyt signifikansnivå er at sjansen for å ha feil øker, men det gjør det enklere å konkludere at koeffisientene er ulike fra null (Hair, Black, Babin, Anderson & Tatham 2006: 216).

En regresjonskoeffisient tilsvarer regresjonsligningens helningsvinkel og brukes for å se på effekten av variabler. Både den ustandardiserte, b , og den standardiserte beta, eller β , vil bli rapportert i de multiple regresjonsanalysene. I bivariat regresjon vil kun beta stå sentralt fordi denne har samme verdi som korrelasjonskoeffisienten Pearsons r . For hver enhet den uavhengige variabelen endrer seg er det en forventet endring tilsvarende størrelsen på den ustandardiserte regresjonskoeffisienten b i avhengig variabel (Pedhazur 1997: 18). Hvis for eksempel *BNP endring* har en b -verdi på 0.3 innebærer det at en vekst i BNP på én prosent gir 0.3 score høyere storhetsgrad. På tross av at b uttrykker retning og størrelse på effekten kan ikke koeffisienten brukes til å sammenligne variabler med ulik måleenhet. Måleenheten til beta er standardavvik og blir oppgitt under antakelse av at alle involverte variabler er z -variabler med standardavvik som måleenhet. I motsetning til b gjør beta det mulig å sammenligne effekter mellom uavhengige variabler når disse er målt på ulike skalaer (Hair *et al* 2006: 224-226).

Standardfeilen til regresjonskoeffisienten, SE_b , forteller oss om gjennomsnittet av helningskoeffisienten er så ulikt null at det kan stamme fra en populasjon der helningen også er ulik null. Av formelen $t=(b-0)/SE_b$ kan dette målet brukes for å finne t -verdien. Med signifikansnivå på 10 prosent følger det at man må ha en t -verdi større enn 1.71 for at variabelen skal være statistisk signifikant (Bohrnstedet & Knoke 1994). Høye standardfeil henger altså sammen med lave t -verdier og tilsvarende høye p -verdier. T -verdiene vil ikke bli gjengitt i analysens modeller, men forholdet mellom b -verdiene og SE_b vil som vist gi en indikasjon på dette nivået. P -verdien, som er et mer presist mål på statistisk signifikans, er den som vil bli benyttet når muligheten for å forkaste nullhypotesene blir diskutert.

Innenfor samfunnsvitenskapen vil det, selv om man inkluderer et stort antall uavhengige variabler, alltid være andre faktorer som også påvirker fenomenet man vil forklare/predikere (Lewis-Beck 1980: 9-10). Residualene estimerer hvor stor del av variasjonen i den avhengige variabelen som er fremkalt av andre prediksjonsvariabler enn dem som er inkludert i analysen (Lewis-Beck 1980: 38). Residualene er således forskjellen mellom faktiske og predikerte verdier målt gjennom formelen $e=Y_f-Y_p$, der e er residual, Y er avhengig variabel og f og p står for henholdsvis faktiske og predikerte verdier. Lave residualverdier innebærer derfor at det er liten forskjell mellom de predikerte verdiene og de faktiske. Begrepene under-og overpredikering kan virke misvisende i så måte. En president blir altså underpredikert hvis de faktiske verdiene er høyere enn de predikerte. Gjennomsnittet av den enkelte statsministers avvik vil alltid være lik null. (Hair *et al* 1998: 153).

Etter at studiens endelige modell er analysert vil residualverdiene bli diskutert. Et sentralt tema her er hvorvidt utelatte variabler kan øke prediksjonen av storhet. Systematiske tendenser i residualene kan gi viktige svar på hvorfor de enkelte statsministerne er vurdert som de er. Det blir satt to kutt punkt for interessante residualverdier, og dette er gjort for å begrense omfanget av analysen. Blant statsministere som har residualverdier over/under 1/-1 vil det diskuteres hva årsakene til dette kan være. Drøftingen vil hovedsakelig være basert på respondentenes utfyllende kommentarer i forbindelse med vurderingene av statsministerne. For regjeringsledere med residualverdier over/under 2/-2 vil det inkluderes en eller flere nye variabler som kan tenkes å være årsak til disse utslagene. Formålet er å analysere hvorvidt disse faktorene kan predikere mer av storhetsgraden til de aktuelle statsministerne enn variablene som allerede er inkluderte.

For å se om enkelte statsministere har stor innflytelse på resultatene blir det kjørt tester på $sdfbeta$ og $sdfit$ ³⁰ (s står for standardiserte verdier). Førstnevnte omhandler enkeltvariabler og predikerer endringen i t-verdien på grunn av ekskluderingen av en spesiell observasjon. Formelen $2/\sqrt{N}$ der N er antall observasjoner gir i dette datasettet en kritisk verdi på 0.4 fordi det inneholder 25 enheter. Statsministere som får verdier over eller under dette nivået på en variabel kan betegnes som innflytelsesrike på denne uavhengige variabelens t-verdi, og disse vil bli analysert nærmere.

³⁰ Se fullstendig oversikt over verdiene for $sdfbeta$ og $sdfit$ i appendiks.

I motsetning til $sdfbeta$, som omfatter regresjonskoeffisienten til hver variabel, påvirker $sdfit$ -verdiene modellens samlede effekt målt gjennom F-verdien. Dette leder oss over på tester og mål som er relevante for modeller. $Sdfit$ gir kritisk verdi gjennom formelen $2\sqrt{p/N}$, der p er antall uavhengige variabler. Målet viser endringen i en predikert variabel på grunn av ekskluderingen av en spesiell observasjon. Den kritiske verdien vil bli nærmere diskutert i analysekapittelet da det foreløpig er uvisst hvor mange variabler som går videre til den endelige modellen.

Gjennom F-testen analyserer man effekten av hele modellen og ikke av hver variabel slik t-verdien gjør. F-testen vil bli brukt for å sammenligne effekten av ulike variabelgrupper, og man kan således si om en modell har større effekt enn de andre. Signifikansen til F-testen forteller oss om modellen som helhet viser en lineær sammenheng som er ulik fra null, og er en indikasjon på om man har signifikante variabler i analysen. Hvis signifikansen er over 0.10 betyr det at modellen ikke klarer å fange opp noen lineær sammenheng mellom de uavhengige variablene og den avhengige.

Forklart varians, R^2 , forteller hvor mye av variasjonene i avhengig variabel Y som kan predikeres av variasjonen i uavhengig variabel X . Justert forklart varians kontrollerer for antall variabler og enheter som er inkludert, og brukes når man har en modell med flere uavhengige variabler (Midtbø 2007: 104-105). En modell med lav forklart varians innebærer at regresjonsligningen har begrenset nytte når det gjelder å predikere enhetenes verdi på den avhengige variabelen (Hellevik 2002: 274). Dette betyr ikke nødvendigvis at modellen er dårlig, men at flere faktorer enn de inkluderte predikerer den uforklarte variansen i avhengig variabel (Skog 2004: 279).

Standardfeilen til regresjonen er et mål på den kvadrerte summen av residualene målt gjennom spredningen rundt regresjonslinjen. Verdien forteller oss hvor dårlig modellen fungerer. Mens R^2 gjør det mulig å avgjøre modellens *relative* prediksjonskapasitet, avgjør standardfeilen til regresjonen modellens *absolutte* prediksjonskapasitet og nøyaktighet fordi den er målt i samme enhet som avhengig variabel. En R^2 -verdi på 0.60 innebærer for eksempel at 60 prosent av variasjonen i avhengig variabel kan predikeres av variasjonen i en eller flere uavhengige variabler. Hvis standardfeilen til regresjonen i denne analysen er på 1 innebærer det derimot at forventet feil i å predikere storhet er på rundt 1 score.

Det er uenighet blant forskere om hvor mye man skal vektlegge R^2 i regresjonsmodeller fordi målet er følsomt for blant annet variasjon i prediksjonsvariabelen, omfanget av målefeil og variasjonen i utelatte prediksjonsvariabler, altså restleddet (Midtbø 2007: 88-89). Hovedargumentet til Krueger og Lewis-Becks er at det beste er å rapportere både justert R^2 og standardfeilen til regresjonen fordi begge målene gir unik og viktig informasjon om prediksjonsmulighetene til modellen (2007).

I denne analysen vil standardfeilen til regresjonen bli brukt for å sammenligne hvor gode modellene er. Dette kan gjøres fordi det kun er én avhengig variabel i alle modellene (Krueger & Lewis-Beck 2007: 3). De ulike funksjonene til standardfeilen til regresjonens og R^2 gjør at begge blir rapportert i analysen.

4.5.2 Prediksjon versus kausalitet

Ifølge Hair et.al. er multivariat regresjonsanalyse gjeldende for to hovedklasser av forskningsformål: prediksjon og kausalitet (1998: 159). Det er viktig å klargjøre forskjellene og likhetene mellom disse to tilnærmingene. Prediksjonsmodeller forteller hvorvidt det er en statistisk sammenheng mellom variabler, men ikke hvordan variabler kan tenkes å påvirke hverandre i kausal forstand. Det gjør derimot kausalmodeller (Hellevik 2002: 290-291). Pedhazur forklarer ulikhetene på denne måten: "In short, explanations provide understanding, but we can predict without being able to understand, and we can understand without necessarily being able to predict" (1997: 196). Det vil derfor ofte være liten forskjell mellom de to analyseformene fordi en prediksjon kan vise seg å ha kausal effekt og visa versa. (Lewis-Beck 1980: 19-20).

Den statistiske fremgangsmåten i de to analysene er i stor grad lik. Ved bruk av regresjon analyserer man effekt/prediksjon av de ulike variablene og variabelgruppene, hvilken retning disse går i og hvorvidt de er signifikante. Forskjellen er i stor grad knyttet til hvor mye man vektlegger ulike mål. Selv om kausalanalyser også fokuserer på residualverdier³¹ kan det argumenteres for at disse verdiene er mer interessante i prediksjonsanalyser. Årsaken til dette er at verdiene viser hvor godt man predikerer hver observasjon. Hovedmålet med denne studien er å analysere hvilke faktorer som predikerer storhet blant norske statsministere. I denne sammenheng kan residualverdiene i analysen kan gi viktige svar på hvorfor noen

³¹ I kausale modeller blir verdiene blant annet brukt til å teste forutsetninger, samt at de danner grunnlaget for forklart varians og standardfeilen til regresjonen.

statsministerprestasjoner blir bedre predikert enn andre og hvilke alternative faktorer som kan predikere storhet.

For at kausalitet skal forekomme må sammenhenger være statistiske. Derfor vil denne analysen uansett resultat være et bidrag til fremtidige studier av årsakssammenhenger. Hovedgrunnen til at prediksjon er viktigere enn kausalitet i studien er at den tar for seg et fagområde det aldri tidligere er gjort tilsvarende analyser på i Norge. Det manglende teorigrunnet gjør således en prediksjonsanalyse mer hensiktsmessig fordi denne konstaterer at det er sammenhenger uten å si at X nødvendigvis er årsaken til Y. Man kan for eksempel slå fast at når arbeidsledigheten er lav så blir storhetsscoren høyere uten at man derfor sier at det er en årsak til storhet. Resultatene fra den kommende analysen kan også føre til usikkerhet knyttet til hvilket kausalt forhold det er mellom de uavhengige variablene og den avhengige. Er det for eksempel slik at dyktige statsministere gjør at det er lav arbeidsledighet, eller er det lav arbeidsledighet som gjør at statsministere blir vurdert som dyktige? Målet med denne studien er å predikere mest mulig av storhetsgraden uten å fokusere på kausale forhold. Derfor vil ikke analysen gå nærmere inn på temaet. Likevel er det et viktig spørsmål som kan være aktuelt å studere i kommende studier av norske statsministeres prestasjoner.

Prediksjon som tilnæringsmåte spiller en relativt liten rolle i statsvitenskapelige studier. Ifølge Krueger og Lewis-Beck (2005:17) er denne neglisjeringen uheldig fordi prediksjon er en viktig dimensjon i vitenskapen, og som metode, for å vurdere alternative generaliseringer eller modeller.

OLS-regresjon gjør at man kan beregne de nevnte residualverdiene, differansen mellom hvert faktiske datapunkt og den estimerte regresjonslinjen (Hellevik 2002: 271). ”Minste kvadratsums metode består i å velge den linjen som gir en minst mulig SSE, eller sum of squared errors” (Skog 1998: 222). Pedhazur forklarer meningen med denne metoden på følgende måte:

”The purpose, then, is to find a solution for the constants, a and b , so that explanation or prediction of Y will be maximized. Stated differently, a solution is sought for a and b so that e – errors committed in using X to explain Y – will be at a minimum” (1997: 19).

I forbindelse med målet om å maksimere prediksjon av storhet brukes flere dummyvariabler. Flere av de uavhengige variablene ble tidligere i kapittelet kategorisert som dummyer ut fra formelen $k-1$ der hver dummyvariabel representerer en kategori, k , av en ikke-metriske uavhengig variabel (Hair *et al* 1998: 167). Fordelen med denne variabeltypen fremfor de som er på nominal- og ordinalnivå er at disse er hevet over måleproblematikken. Man står således fritt til å anvende statistiske teknikker uansett hvilket målenivå disse forutsetter (Hellevik 2002: 182-183; Midtbø 2007).

4.6 Måle- og spesifikasjonsfeil

Målefeil representerer den graden de observerte verdiene ikke er representative for de ekte verdiene. Dette er ofte et problem i multivariate analyser (Hair *et al* 1998: 9). ”Vi har å gjøre med målefeil når den observerte variasjonen skyldes egenskaper med måleinstrumentene og ikke fenomenet selv” (Midtbø 2007: 25). Målefeil kan knyttes til begrepene reliabilitet og validitet. Selv om disse ikke nødvendigvis er årsaken til dårlige resultater, vil tilstedeværelsen av dem alltid forstyrre de observerte sammenhengene og gjøre multivariate teknikker svakere (Hair *et al* 1998: 10).

Reliabilitet avhenger av hvordan datainnsamlingen er foretatt og sikter til målingens nøyaktighet. Lite reliable data minker sjansen for å oppnå likt resultat hvis man gjennomfører samme studie flere ganger. Målet på arbeidsledighet i denne studien kunne hatt et potensielt reliabilitetsproblem fordi forutsetningene for det innsamlede tallmaterialet har endret seg gjennom perioden 1905-1997. Spesielt i perioden før Andre verdenskrig finnes det ulike mål på arbeidsledigheten. Dermed kan man stille spørsmål til om målingene

er nøyaktige nok (Midtbø 2007: 25). I denne studien er likevel arbeidsledighetsdataene fra perioden før 1960 hentet fra en fagspesialist som har hatt fokus på å gjøre tallene mest mulig sammenlignbare med tall etter denne perioden, og dette styrker reliabiliteten (Grytten 1994).

Validitet dreier seg om hvorvidt vi måler det vi faktisk ønsker å måle. ”Validitetsproblemet oppstår når vi ikke kan måle den egenskapen vi har definert teoretisk direkte” (Hellevik 2002: 187). For å sikre validiteten må man ha en forståelse av hva som skal måles og deretter måle dette så korrekt og nøyaktig som mulig (Hair *et al* 1998: 9). Innenfor samfunnsvitenskapen og kvantitativ metode spesielt er validitetsproblemet vanlig fordi man som oftest må måle effekter gjennom variabler. Man kan for eksempel ikke måle sosioøkonomiske forhold direkte, men ved å operasjonalisere begrepet gjennom BNP og arbeidsledighet er det mulig å gjøre dette indirekte. Et eksempel på mulig validitetsproblem i denne analysen knyttes til variabelen *gjenvalg*. Statsministeren Ivar Lykke ble per definisjon *gjenvalt* fordi han fortsatte å regjere etter stortingsvalget i 1927. Likevel gjorde Høyre et dårlig valg og etter tre måneder med mindretallsregjering tok Christopher Hornsruds Arbeiderpartiregjering over. Dette var på grunn av Høyres dårlige valgresultat, og det er tvilsomt om Lykkes *gjenvalg* har positiv prediksjon på storhet. Dette svekker variabelens validitet. I tillegg har flere statsministere både blitt ”*gjenvalt*” og mistet regjeringsmakt etter stortingsvalg. I slike tilfeller svekkes også validiteten.

Spesifikasjonsfeil innebærer at irrelevante variabler inkluderes i analysen mens relevante variabler utelates. Dette kan blant annet føre til at signifikanstesting av uavhengige variabler blir mindre presis, samt at den statistiske og praktiske betydningen av analysen reduseres. På tross av at variablene i denne analysen er inspirert av internasjonal forskning og norske historikers betraktninger, gjør det manglende forskningsgrunnlaget i Norge at spesifikasjonsfeil lettere kan oppstå enn innenfor studier der teorigrunnlaget er bredere. Likevel er dette en utfordring som er vanskelig å unngå innenfor nye forskningsfelt.

4.7 Kriterier i analyseprosessen

Fremgangsmåten for de kommende analysene er at institusjonelle, sosioøkonomiske og biografiske variabler blir testet i tre separate modeller. Faktorer som er relevante for prediksjon av storhet vil deretter bli inkludert i en endelig modell. Det er resultatene fra denne som hovedsakelig vil være analysens bidrag til forskningen. Hvilke variabler som er relevante fra de tre variabelgruppene blir bestemt ut fra følgende kriterier: hvorvidt

variablene 1) er substansielt signifikante. Dette innebærer at de må være valide og gi mening i konteksten. 2) er statistisk signifikante, og 3) har høye toleranseverdier. Det vil si at det er høy uforklart varians i den uavhengige variabelen.

Det vil bli gjennomført flere prosedyrer for å sjekke kriteriene. Dette kapittelets diskusjon om operasjonaliseringen av variabler ga en indikasjon på hvilke faktorer som ble regnet for å ha mindre substansiell signifikans. For eksempel har *gjenvalg* et potensielt validitetsproblem fordi flere statsministere både har blitt gjenvalg og ”tapt” valg. Lignende problemer kan knyttes til blant annet *leder av flertallsregjering* og *koalisjonsregjering*. *Arbeidsledighet* er derimot en mer substansielt signifikant prediksjonsfaktor fordi variabelen har relativt høy validitet og er målt på kontinuerlig skala. I forbindelse med substansiell signifikans mener Gujarati at ”the researcher should be more concerned about the logical or theoretical relevance of the explanatory variables to the dependent variable and their statistical significance (1995: 21). Man bør altså ikke utelukkende fokusere på den statistiske signifikansen.

Denne studien er i stor grad utforskende, og det vil derfor være galt å forkaste variabler før de er testet statistisk. Derfor vil også prediksjonsfaktorer som har antatt lav validitet bli analysert. Det vil herfra diskuteres hvilke variablene som skal inkluderes i den endelige modellen.

Statistisk signifikans vil bli testet for i bivariat- og multivariat regresjon. Det vil også bli kjørt test på *backward regression*. Dette er en statistisk elimineringsmetode der man starter med en regresjonsligning som inkluderer samtlige 21 uavhengige variabler. Deretter elimineres variabler som ikke bidrar signifikant til prediksjonen av storhet. Metoden indikerer således hvilke prediksjonsvariabler som er sentrale i analysen når det legges vekt på effektivitet og styrke (parsimony and power). Hvis substansielt sentrale faktorer viser seg å være signifikante i *backward regression* vil de også kunne bli inkludert i den endelige modellen.

Høye toleranseverdier vil bli testet for i de gruppevise multivariate regresjonene og i den endelige modellen. Det er viktig å presisere at de nevnte kriteriene ikke vil være absolutte. Det vil foretas en helhetsvurdering av prediksjonsfaktorene og på grunnlag av denne blir det bestemt hvorvidt den enkelte variabel er god nok til å bli tatt med videre i analysen. For

eksempel vil substansielt signifikante verdier som er signifikante i bivariat regresjon, men ikke i multivariat, kunne inkluderes i modellen. Valget følger argumentet til Gelman og Stern om at små forskjeller i signifikansnivået ofte kan ha sammenheng med små insignifikante endringer i de underliggende dataene (2006: 328). Differansen mellom signifikante og insignifikante verdier er altså ikke nødvendigvis signifikant. Derfor skal man være åpen for å inkludere insignifikante variabler fra multivariat regresjon. Også prediksjonsfaktorer som bidrar signifikant til prediksjon av storhet i *backward regression* vil kunne inkluderes i den endelige modellen.

4.8 utfordringer knyttet til tid og OLS-regresjon

Det opereres vanligvis med tre analyseformer som skiller seg fra hverandre ved at de fokuserer på enten tid, rom eller begge deler. Tverrsnittanalyse innebærer at man samler inn data på ett enkelt tidspunkt mens tidsserieanalyse bruker data som endrer seg over tid. Panelanalysen er en kombinasjon av tid og rom.

I utgangspunktet kan denne studien av 25 statsministere i perioden 1905-1997 betegnes som en tverrsnittanalyse fordi hver person er grunnlag for én observasjon. Likevel har den elementer av tid i seg. For eksempel er *BNP endring* målt over flere år, og dermed støter man på noen av de samme utfordringene som i strenge tidsseriestudier.

Autokorrelasjon i restleddet er ett potensielt problem og innebærer at en variabels verdier er korrelert med tidligere verdier av seg selv. Det betyr at lave verdier etterfølges av nye lave verdier og visa versa. Dette kan for eksempel vise seg å være tilfellet med arbeidsledighet fordi endringene vanligvis utvikler seg over tid. I analyser som har elementer av tid i seg pleier "[v]erdier på ett tidspunkt å være avhengige av verdier på tidligere tidspunkt, og modellering av denne avhengigheten er tidsserieanalysens største utfordring" (Midtbø 2007: 116). I denne analysen vil dette bli testet for ved hjelp av Durbin-Watson. Målet går fra 0 til 4 hvor 0 innebærer fullstendig positiv autokorrelasjon og 4 fullstendig negativ. Verdier mellom 1,5-2,5 indikerer at autokorrelasjon ikke er et problem. Ifølge Chatfield er målet viktig for å beskrive en tidsserie og variabelnes avhengighet over tid (1996: 18-19 og 30).

Multikollinearitet er en annen utfordring med tidsserier og oppstår hvis uavhengige variabler med høy innbyrdes korrelasjon inkluderes i samme multivariate regresjonsanalyse. Multikollinearitet representerer den graden en variabels effekt blir predikert av andre

uavhengige variabler i analysen (Hair *et al* 1998: 24). Det finnes ulike meninger om hvor høy korrelasjonen kan være før dette blir et problem. Ifølge Lewis-Beck (1980: 60) er verdier rundt 0.8 eller høyere problematiske mens Hair *et al.* (Hair *et al* 1998: 191) opererer med 0.9. Denne analysen vil benytte seg av 0.8 som kritisk verdi.

Resultatet av multikollinearitet er at variablenes samlede varians øker mens hver variabels unike varians minker. Dette fører til at det knyttes økt usikkerheten til regresjonskoeffisientene. I ikke-eksperimentelle samfunnsvitenskapelige data er dette et svært vanlig problem (Hellevik 2002: 289; Lewis-Beck 1980). Hair *et. al* mener man for å maksimere prediksjonsgraden fra et gitt antall uavhengige variabler bør inkludere de som har lav innbyrdes korrelasjon og høy korrelasjon med avhengig variabel (2006: 186). Det finnes ifølge Lewis-Beck flere symptomer på multikollinearitet:

”One symptom is a substantial R^2 for the equation, but statistically insignificant coefficients. Another, but weaker, is regression coefficients which change greatly in value when independent variables are dropped or added to the equation” (1980: 60).

Multikollinearitet vil først bli testet mellom de enkelte uavhengige variablene i en korrelasjonsmatrise og deretter ved å se på toleranseverdier i de multivariate regresjonsanalysene. Toleranse er definert som én minus den forklarte variansen i en uavhengig variabel, og innebærer variasjonen i den uavhengige variabelen som ikke er forklart av de andre uavhengige variablene i regresjonslikningen³² (Hair *et al* 2006: 227). I denne analysen bør variabler ha toleranseverdier over 0.6 for å kunne inkluderes i den endelige modellen.

4.9 Oppsummering av data, operasjonalisering og metode

Dette kapitlet består av tre hoveddeler. Den første omhandler problemstillinger knyttet til dataene som ligger til grunn for denne analysens avhengige variabel. Det ble foretatt en gjennomgang av Cordtsen-studien der fremgangsmåten for undersøkelsen, og utfordringer knyttet til denne, ble diskutert. Konklusjonen var at statsministernes storhetsscore i Cordtsen-studien samsvarte med utregningene i denne analysen.

³² For tester av regresjonsanalysens forutsetninger se appendiks E.

Kapittelet drøftet deretter hvorvidt respondentene i denne type undersøkelser er objektive i sine vurderinger. Dette temaet har blitt diskutert i den internasjonale litteraturen og er spesielt aktuelt i studier der respondentene selv bestemmer kriteriene for storhetsgrad (se for eksempel Bose & Landis 2003; Simonton 1986; Theakston & Gill 2006). I denne sammenheng ble fokuset rettet mot fire potensielle utfordringer: 1) Påvirker respondentenes politiske ståsted deres vurderinger av statsministerprestasjoner? Norske akademikere er generelt mer politisk venstreorienterte enn den øvrige befolkningen (Aardal *et al* 2007: 13). Murray og Blessing mener dette kan påvirke deres bedømmelser (1983: 543). Cohen er uenig i dette og viser til analyser som tyder på at respondentene vurderer uavhengig av politisk ståsted (2004: 2-3). 2) Har respondentene egne hensikter i forbindelse med vurderingene? Det ble her nevnt at Arthur M.Schlesinger jr. har vært politisk rådgiver for John F.Kennedy. Selv om han selv ikke har tatt del i rangeringene kan slike roller svekke respondentenes troverdighet. 3) Spiller kjønn en rolle i statsministervurderinger? Alle respondentene i Cordtsten-undersøkelsen er menn. Ifølge Bose og Landis er dette et problem fordi mannlige respondenter bedømmer regjeringslederens prestasjoner ut fra andre kriterier enn kvinner (Bose & Landis 2003: 93). 4) Blir respondentene påvirket av hvordan statsministerne er blitt omtalt gjennom historien og i media? Dette er en potensiell svakhet som sannsynligvis øker jo lengre tilbake i tid statsministeren regjerte. Crockett mener at dette er et problem i studier av presidenters prestasjoner (2002: 239-241).

På tross av disse mulige utfordringene knyttet til respondentenes objektivitet viser Murray og Blessing til studier som har konkludert med at demografiske faktorer i svært liten grad påvirker respondentenes rangeringer (1983). I tillegg er det i USA funnet høy korrelasjon mellom respondenters vurderinger på tvers av studier (Simonton 1986: 262). Det ble også argumentert for at man må forvente at historikerne og samfunnsviterne er profesjonelle og ikke lar vurderingene bli påvirket av deres personlige meninger. Likevel er det mulig at de nevnte utfordringene påvirker verdiene på den avhengige variabelen. Hvorvidt dette er tilfellet er vanskelig å gi svar på. Formålet med studien er prediksjon av storhet, ikke inngående analyser av respondenter.

I kapittelets andre hoveddel ble operasjonaliseringen av de institusjonelle, sosioøkonomiske og biografiske prediksjonsvariablene diskutert. Formålet med denne delen var å identifisere mulige validitetsproblemer knyttet til dataene og kodingen av dem, samt forsvare de valgene som ble gjort.

Del tre tok for seg metodiske temaer knyttet til den kommende analysen. Etter å ha forsvart valget av kvantitativ metode ble det fokusert på multivariat regresjon, der sentrale mål og tester ble gjennomgått. Det ble blant annet foretatt et valg om å rapportere både R^2 og standardfeilen til regresjonen i analysen fordi begge målene gir verdifull og unik informasjon om modellenes forklaringskraft.

I denne studien er prediksjon foretrukket fremfor kausalitet. Årsaken til dette er at analysen i stor grad er utforskende. Det knyttes derfor usikkerhet til det teoretiske grunnlaget for valg av variabler. Således er det mer hensiktsmessig å påvise at det er statistiske sammenhenger mellom de uavhengige variablene og den avhengige enn å konstantere at det er kausale effekter.

Problemer knyttet til reliabilitet og validitet er aktuelle i multivariate analyser. I kapittelet ble disse utfordringene drøftet med eksempler. Diskusjonen førte videre til presisering av følgende kriterier som skal legges til grunn for relevante variabler i analysen: 1) substansiell signifikans, 2) statistisk signifikans og 3) høye toleranseverdier. Kriteriene ikke er absolutte, men heller et utgangspunkt for utvelgelsen av relevante prediksjonsfaktorer.

Til slutt fokuserte dette kapittelet på utfordringer relatert til tid og OLS-regresjon, og da særlig autokorrelasjon og multikollinearitet.

Målet med kapittelet har vært å forsvare valg som er gjort i forbindelse med variabler og operasjonaliseringen av dem, samt valg av metode. Kriterier for hvilke variabler som er sentrale i analysen var også en nødvendig gjennomgang, for å kunne foreta analyser av storhet blant norske statsministere.

5. ANALYSE OG RESULTATER

Forskningsspørsmålet denne oppgaven skal prøve å besvare er: *Hvilke faktorer predikerer graden av storhet blant norske statsministere i perioden 1905-1997?* Det ble i kapittel 3 argumentert for at de inkluderte faktorene predikerer statsministernes prestasjoner, og i forrige kapittel ble kodingsvalgene diskutert. I denne delen det foretas statistiske tester av variablene.

5.1 Struktur

Dette kapitlet består av to hoveddeler. I del 1 blir det først presentert deskriptiv statistikk for å vise oversikten over fordelingsegenskapene til samtlige variabler. Videre vil korrelasjon mellom de uavhengige variablene bli diskutert for å kontrollere at det ikke inkluderes overflødige variabler i analysen. Regresjonsanalysen gjennomføres ved bivariate analyser av variablenes prediksjonseffekt. Dette vil gi en første indikasjon på hvilke faktorer som er relevante for statsministernes prestasjoner. Deretter blir det kjørt separate multivariate regresjonsanalyser av de *institusjonelle*, *sosioøkonomiske* og *biografiske* variabelgruppene. Eventuelle signifikante variabler uten toleranseverdier blir deretter analysert i en endelig modell. I forkant av denne modellen blir det også kjørt test på *backward regression*. Kriteriene fra kapittel 4 danner grunnlaget for en helhetlig vurdering av hvilke variabler som skal analyseres videre.

I del 2 blir residualene fra den endelige modellen analysert, og disse vil vise hvilke statsministere som blir godt og dårlig predikert av modellen. Deretter følger en diskusjon av årsaker til avvik i residualene. Det blir også testet om enkelte statsministere har stor innflytelse på resultatene i den endelige modellen. Prediksjon av storhetsgraden til statsministere som har regjert etter 1997 blir til slutt beregnet med utgangspunkt i denne modellen.

5.2 Variablenes fordeling

Ved å se på deskriptiv statistikk kan man få bedre kjennskap til hvordan variablenes verdier fordeler seg. Den laveste og høyeste variabelverdien, gjennomsnittet og standardavviket er de målene blir diskutert her. For *utdanning* er variabelverdiene: 0) *grunnskole*, 1) *over grunnskole* og 2) *hovedfag*. For *region* er variabelverdiene: 0) *region Nord-Norge*, 1) *region Sør-Vest Norge* og 2) *region Østlandet*.

Tabell 2 – Deskriptiv statistikk av variablene i analysen

Variabel	Minimum	Maksimum	Gj.snitt	St.avvik
Storhet	1.45	5.85	3.29	1.29
Annet parti enn sin forgjenger	0	1	0.72	0.46
Personlige årsaker hos forgjenger	0	1	0.20	0.41
Leder av koalisjonsregjering	0	1	0.48	0.51
Leder av flertallsregjering	0	1	0.36	0.49
Antall måneder	1	209	45.40	50.84
Flere statsministerperioder	0	1	0.28	0.46
Partipolitisk skillelinje	0	1	0,60	0.50
Gjenvalg	0	1	0,48	0.51
Avgang pga konfrontasjon	0	1	0,44	0.51
Avgang pga intern stridighet	0	1	0,24	0.44
Avgang pga valg	0	1	0.16	0.37
BNP endring	-1.10	5.40	2.59	1.47
Arbeidsledighet	0.60	10.40	3.58	3.24
Utdanning	0	2	1.28	0.74
Region	0	2	1.48	0.77
Alder ved innsettelse	41	74	55.44	8.02
Leveår etter avsettelse	0	33	17.08	10.20

N=25

Tabellen viser at verdiene på den avhengige variabelen *storhet* varierer relativt mye mellom statsministerne, fra Jens Hunsheids minimumsverdi på 1.45 til Einar Gerhardsens maksimumsverdi, 5.85. Gjennomsnittsscoren til statsministerne i perioden er 3.29 og standardavviket er 1.29. Sistnevnte innebærer at respondentenes vurdering av statsministerne avviker 1.29 score fra gjennomsnittet. Tabellen viser også stor variasjon i *antall måneder*. Mens Christopher Hornsrud og John Lyng satt én måned hver var Einar Gerhardsen regjeringsleder i til sammen 209 måneder. Statsministerne i perioden satt gjennomsnittlig 45 måneder med makten. Det er også relativt store ulikheter når det gjelder statsministernes verdier på den sosioøkonomiske variabelen *arbeidsledighet*. I gruppen av biografiske variabler ser man at *alder ved innsettelse* varierer mellom 41 og 74 år med et gjennomsnitt på 55. Foruten dette lever statsministerne i den aktuelle perioden gjennomsnittlig 17 år etter de ble avsatt.

Variablene *antall måneder* og *arbeidsledighet* har begge høye standardavvik. Dette innebærer at enkelte statsministere trekker gjennomsnittet kraftig opp, og fører til skeivfordelte variabler. *Antall måneder* har større standardavvik enn variabelens snitt, hovedsakelig på grunn av Gerhardsens 209 måneder lange regjeringstid. Lignende blir

arbeidsledighet skeivfordelt, særlig på grunn av den internasjonale nedgangstiden som rammet landet i mellomkrigsårene. Perioden toppet seg med over ti prosent ledighet under Jens Hundseids regjeringstid.

Alternativet for å minke skeivfordelingen er å omskalere *antall måneder* og *arbeidsledighet*. Dette vil imidlertid skape problemer fordi variablene da mister sin måleenhet. Variabelverdiene varierer svært mye mellom observasjonene, og en omskalering vil gjøre målet mer upresist og generelt enn utgangspunktet. Ved å ha intervaller der flere statsministere havner innenfor samme kategori, vil man for eksempel ikke kunne nyansere forskjellene fordi man ikke får frem alle variasjoner i variabelverdiene. Beslutningen om å la variablene være i sin originale form er tatt ut fra en avveining om hvorvidt man skal fokusere på det metodisk mest korrekte eller det substansielt viktige, der sistnevnte ble foretrukket. Valget gjør det også lettere å sette inn verdier for prediksjon, noe som er aktuelt i forhold til statsministerne etter 1997.

5.3 Regresjonsanalyser

I forkant av de kommende analysene er det foretatt korrelasjonsanalyser mellom samtlige uavhengige variabler (se appendiks 3). Ingen har Pearsons r-verdier over 0.7, og dette innebærer at de i stor grad måler ulike variasjoner i storhet. Alle variablene kan derfor inkluderes i regresjonsanalysene.

Forholdet mellom variablene innenfor de institusjonelle, sosioøkonomiske og biografiske blokkene og storhetsgrad vil nå bli analysert, først ved bivariat regresjon og deretter ved multiple regresjonsanalyser.

5.3.1 Bivariat regresjon

Den bivariat regresjonen i tabell 3 viser sammenhengen mellom hver uavhengige variabel og den avhengige. Hensikten med bivariat regresjon er å få et inntrykk av hvilke variabler som vil være vesentlige i de kommende analysene.

Tabell 3 – bivariat regresjon

uavhengige variabler	grad av storhet beta / p- verdi
Antall måneder	0.711 / 0.000
Flere statsministerperioder	0.557 / 0.004
Annet parti enn sin forgjenger	0.400 / 0.850
Gjenvalg	0.361 / 0.076
Arbeidsledighet	-0.352 / 0.084
Utdanning over grunnskole	-0.325 / 0.113
Leder av flertallsregjering	0.316 / 0.124
BNP endring	0.305 / 0.138
Innsatt pga. personlige årsaker hos forgjenger	-0.253 / 0.223
Alder ved innsettelse	-0.238 / 0.253
Avgang pga. Valg	-0.184 / 0.378
Avgang pga. Konfrontasjon	0.160 / 0.444
Leder av koalisjonsregjering	-0.156 / 0.457
Utdanning grunnskole	0.121 / 0.564
Partipolitisk skillelinje	0.107 / 0.612
Region Nord-Norge	0.106 / 0.613
Leveår etter avsettelse	-0.034 / 0.873
Avgang pga. intern strid	0.031 / 0.884
Region Sør-Vest Norge	-0.020 / 0.924

Tabell 3 er sortert etter verdi på beta-koeffisienten, og man ser at lave signifikansverdier korresponderer med høye betaverdier. Dette vil si at variablenes effekt på storhet kan knyttes til hvor sannsynlig det er at sammenhengene ikke er tilfeldige. *Annet parti enn sin forgjenger* skiller seg i så måte ut fra de andre uavhengige variablene. Den høye p-verdien kombinert med høy betaverdi gjør at det knyttes mer usikkerhet til den reelle effekten av variabelen, og det blir derfor interessant å se om p-verdien minker når det i regresjonsanalysene kontrolleres for andre faktorer.

Tabellen viser ellers at følgende fire variabler har signifikante verdier: *antall måneder*, *flere statsministerperioder*, *gjenvalg* og *arbeidsledighet*. Det er altså sannsynlig at disse sammenhengene er reelle. Lengden på statsministerperioden har sterkest sammenheng, 0.711, med endringer i storhetsscoren. Dette er en indikasjon på at variabelen vil være sentral i prediksjonen. Ellers antyder tabellen at det i stor grad er de institusjonelle og

sosioøkonomiske variablene som har størst betydning. Tre av de fire prediksjonsfaktorene med lavest betaverdi i bivariat regresjon tilhører gruppen av biografiske variabler, noe som ikke er overraskende. I den internasjonale litteraturen blir slike faktorer vurdert som marginale i prediksjon av regjeringslederens storhetsgrad (Murray & Blessing 1994: 25-32). Det er viktig å presisere at selv om den bivariante analysen indikerer variabelenes relevans, så kan både prediksjonens retning og effekt endres når det kontrolleres for andre uavhengige variabler i de multivariate regresjonsanalysene.

5.3.2 Institusjonelle faktorer

Elleve variabler utgjør den institusjonelle gruppen vist i tabell 4. Faktorene er relatert til hvorfor statsministerne ble innsatt og avsatt, samt regjeringstype og tidsepoken de regjerte i. Generelt anses variabelgruppen som substansielt viktigere enn de sosioøkonomiske og biografiske. Den bivariante analysen (tabell 3) indikerte at *antall måneder* vil være sentral i prediksjon av storhet. Det knyttes derimot mer usikkerhet til *annet parti enn sin forgjenger*.

Tabell 4 – institusjonell modell

Variabel	b	St.feil til reg.koeff.	beta	p- verdi	toleranse
Konstant	1.411	1.077		0.213	
Annet parti enn sin forgjenger	0.589	0.716	0.209	0.425	0.424
Personlige årsaker hos forgjenger	-0.325	0.756	-0.103	0.674	0.480
Leder av koalisjonsregjering	0.075	0.603	0.030	0.902	0.484
Leder av flertallsregjering	0.465	0.655	0.176	0.491	0.443
Antall måneder	0.018	0.008	0.715	0.035	0.297
Flere statsministerperioder	0.309	0.702	0.110	0.667	0.442
Partipolitisk skillelinje	0.441	0.546	0.171	0.433	0.613
Gjenvalg	-0.334	0.626	-0.132	0.602	0.448
Avgang pga konfrontasjon	0.463	0.817	0.182	0.581	0.267
Avgang pga intern stridighet	0.242	0.869	0.082	0.785	0.318
Avgang pga valg	0.244	0.976	0.071	0.807	0.342

R²=0.644 justert R²=0.343 F=2.137 F.sign.=0.097 DW=1.497 St.feil til regresjonen=1.047

Modellens justert R² er 34.3 prosent. Dette innebærer at over en tredel av variasjonen i grad av storhet blant norske statsministere i perioden kan predikeres av disse faktorene.

Standardfeilen til regresjonen er 1.047, men denne verdien blir først interessant når man sammenligner med de andre gruppene av uavhengige variabler. Da kan man se hvilke modeller som predikerer storhet best.

F-testen gir verdien 2.137 med p-verdi på 0.097. Modellens samlede effekt er med det signifikant på 10 prosent nivå, og modellen har således en lineær sammenheng med grad av storhet.

DW-verdien er på 1.497. Det er lavere enn det kritiske punktet som er satt til 1.5 i denne analysen, og tyder på autokorrelasjon. Foreløpig er ikke dette et problem da det forventes at få av disse variablene går videre til den endelige modellen.

Toleranseverdiene viser at multikollinearitet er tilstede i modellen. Årsaken er at flere variabler måler noe av den samme prediksjonseffekten på storhet. For eksempel er det sannsynlig at *gjenvalg*, *antall måneder* og *flere statsministerperioder* har overlappende varians. Dette er imidlertid heller ikke et problem på dette stadiet av analysen da kun en av variablene er signifikant.

En sammenligning av verdiene for standardfeilen til regresjonskoeffisientene (SEb) og b-verdiene viser at det kun er *antall måneder* som overstiger det kritiske 1.71-nivået for t-verdiene (b/SEb). Samtlige av de andre variablene har høyere verdier på SEb enn b. *Antall måneder* er også den eneste variabelen som er signifikant på ti prosent nivå og som foreløpig går videre til den endelige modellen. Verdien på den ustandardiserte koeffisienten b er 0.018. Dette betyr isolert sett at én måned ekstra regjeringstid predikerer 0.018 høyere storhetsscore. Det er også interessant at kun denne variabelen får sterkere prediksjonseffekt målt gjennom standardisert regresjonskoeffisient, 0.715, enn ustandardisert. Dette innebærer at variabelen styrker seg når det blir kontrollert for andre institusjonelle faktorer, noe som øker sjansen for at det er en viktig faktor.

Det at lengden en statsminister sitter med makten har signifikant positiv prediksjon på storhet er i samsvar med oppgavens hypotese og resultatene fra den internasjonale litteraturen. I Simontons studier i USA (1986: 286-287) er *antall år som president* sammen med *antall år i krig og skandaler* de viktigste predikatorene på storhetsgrad. I canadiske storhetsstudier står lengden på statsministerperioden også sentralt (Ballard & Suedfeld 1988: 297-300). I Theakston og Gills analyse av britiske statsministers grad av storhet hadde ingen av de ti statsministerne med høyest score sittet med makten mindre enn fem år, mens kun to av de ti nederste i rangeringen hadde regjert mer enn fire år (2006: 205). I denne studien har de fire statsministerne som har regjert lengst kommet blant de fem med høyest

score. Unntaket er Michelsen som trolig får høy score på grunn av unionsoppløsningen. Blant de fem statsministerne med dårligst score har ingen regjert lenger enn 25 måneder, og tre av dem har ikke styrt lenger enn ett år.

Tabellen viser at variabler knyttet til regjeringskonstellasjoner og utskiftninger har liten prediksjon på grad av storhet. *Annet parti enn sin forgjenger* har den høyeste b-verdien av samtlige variabler, men den mister effekt når det kontrolleres for de andre institusjonelle faktorene. Variabelen er ikke signifikant, men får lavere signifikansverdi enn 0.850 som var tilfellet innenfor bivariat regresjon. Som det ble forventet i hypotesekapittelet har variabelen positiv effekt på storhetsgrad.

Innsatt pga. personlige årsaker hos forgjenger har b-verdi på -0.325 og betaverdi på -0.103 . Den negative prediksjonen er som forventet. I hypotesekapittelet ble det antatt at når en statsminister blir innsatt som følge av *personlige årsaker hos forgjenger*, skjer slike overganger gjerne midt i en regjeringsperiode. Denne type uforutsette endringer gjør det vanskeligere å komme inn og gjøre en god statsministerjobb enn hvis personen blir innsatt etter et stortingsvalg. Variabelen er likevel insignifikant, og blir foreløpig ekskludert fra videre analyse.

Leder for koalisjonsregjering har positiv, om enn ikke signifikant, prediksjon på avhengig variabel. Dette er overraskende da det ble antatt at slike regjeringer er ustabile og bygger på kompromisser mellom partier, noe som fører til minkende handlingskraft. I hypotesekapittelet ble det likevel diskutert en alternativ hypotese som sa at hvis en statsminister klarer å føre en stabil politikk og få gjennomslag for saker under disse omstendighetene, kan det bli vurdert positivt i ettertid. Dette kan være årsaken til variabelens positive prediksjon.

Leder for flertallsregjering har som forventet positiv prediksjon. Det ble antatt at slike regjeringer skaper stabilitet og gjør det enklere for statsministeren å vise handlingskraft. P-verdien på 0.491 er likevel langt fra signifikant. Dette er overraskende, spesielt siden parlamentarisk flertall i stor grad ble vektlagt av respondentene i Storbritannia (Theakston & Gill 2006: 207).

Det ble i hypotesekapittelet antatt at statsministerne får predikert høyere score for å regjere i flere ulike perioder, noe som også bekreftes i modellen. *Flere statsministerperioder* har likevel svakere prediksjon enn antatt og høy p-verdi, 0.667.

Partipolitisk skillelinje har positiv prediksjon og høy p-verdi. Retningen er ikke som forventet da det i hypotesekapittelet ble antatt at det politiske landskapet ble mer uoversiktlig etter 1927. Nye skiller mellom venstre- og høyresiden oppsto, og flere partier ble etter hvert dannet. Dette førte trolig til mindre makt rundt statsministerrollen enn tidligere. Likevel ble det i hypotesekapittelet knyttet usikkerhet både til variabelens relevans og prediksjon, og det er således ikke overraskende at den er insignifikant.

Gjenvalg har negativ prediksjon, noe som er svært overraskende da det ble antatt at en statsminister som ble gjenvalgt gjorde kloke valg. Personen fikk i tillegg lengre tid til å påvirke politikken og få gjennomslag for regjeringens saker. B-verdien er på -0.334 mens betaverdien ligger på -0.132 . Gjenvalg fører altså til 0.334 prosent lavere grad av storhet, men variabelen mister relativt mye prediksjonskraft når det kontrolleres for andre institusjonelle faktorer. Ifølge Simonton er gjenvalg en viktig predikator på storhetsgraden til amerikanske presidenter (1986: 268). Den ble i tillegg fremhevet som viktig av britiske respondenter i vurderingen av landes statsministere (Theakston & Gill 2006: 205). Den negative prediksjonen skyldes trolig relativt høy korrelasjon, 0.624, med *antall måneder*. Når den overlappende variansen fordeles på avhengig variabel, kan det føre til at prediksjonseffekten til enkeltvariabler endrer seg i statistiske modeller. Ellers ble Otto Blehr, som har lav storhetsscore, gjenvalgt i mellomkrigstiden uten å sitte lenge med makten. Han ble dessuten innsatt kort tid før et stortingsvalg³³, og slike omstendigheter trenger ikke nødvendigvis å gjøre at gjenvalg har positiv prediksjon på statsministerprestasjoner.

De tre variablene som gikk på statsministerens avgang ble i hypotesekapittelet antatt å ha negativ effekt, men det ble knyttet alternative hypoteser til samtlige av dem fordi avganger trolig også kan tolkes positivt av respondentene.

³³ Otto Blehr kom til makten 22.juni 1921 og ble gjenvalgt oktober samme år.

Avgang pga. konfrontasjon har positiv prediksjonseffekt i modellen. Dette kan bety at statsministerne som ble vurdert kontroversielle i sin tid, og ikke vek unna for sine verdier, har blitt vurdert positivt på lengre sikt.

Avgang pga. intern stridighet har, noe overraskende, positiv prediksjon på storhet. Den alternative hypotesen til variabelen bygget på de samme argumentene som *konfrontasjon*. Kontroversielle og prinsippfaste statsministere kan ha blitt vurdert mer positivt i ettertid enn da de styrte.

Avgang pga. valg har også positiv prediksjonseffekt, og årsaken til dette ligger trolig i de samme faktorene som ble nevnt angående de alternative hypotesene til *konfrontasjon* og *intern strid*.

Analysen av de elleve institusjonelle variablene viste at kun *antall måneder* var signifikant. I forhold til den internasjonale litteraturen var det forventet at denne faktoren ville være sentral i prediksjonen av storhet. Det var således mer overraskende at *gjenvalg* var sterk insignifikant. Resultatene fra denne variabelgruppen gjør at *antall måneder* er den eneste variabelen som foreløpig blir inkludert i den endelige modellen.

5.3.3 Sosioøkonomiske effekter

Variabelgruppens to variabler er mål på makroøkonomiske forhold i Norge. Med utgangspunkt i den bivariante analysen er det mest sannsynlig at *arbeidsledighet* vil ha signifikant verdi i modellen vist i tabell 5.

Tabell 5 – sosioøkonomisk faktorer

Variabel	b	St.feil til reg.koeff	beta	p- verdi	toleranse
Konstant	3.147	0.602		0.000	
BNP endring	0.229	0.170	0.261	0.191	0.981
Arbeidsledighet	-0.126	0.077	-0.316	0.117	0.981

R²=0.191 justert R²=0.118 F=2.599 F.sign=0.097 DW=2.052 St.feil til regresjonen=1.213

Modellen med de sosioøkonomiske variablene har en forklart varians på 19.1 prosent mens justert forklart varians er 11.8 prosent. Standardfeilen til regresjonen er 1.213, altså litt høyere enn hva tilfellet var i den institusjonelle variabelgruppen. Verdien innebærer at modellen ikke er like god til å predikere storhetsgraden som den foregående. Dette har

sammenheng med at det var mange flere institusjonelle variabler, noe som trolig har redusert den uforklarte variansen.

F-testen gir verdien 2.599. Effekten er med andre ord sterkere sammenlignet med den institusjonelle modellen. Den signifikante verdien tyder på at modellens samlede effekt er reell. Makroøkonomiske forhold har altså betydning for hvordan respondentene vurderer norske statsministers prestasjoner. Dette er som forventet ettersom historikere vektlegger variablene i sine beskrivelser og analyser av regjeringsledere. Ser man sosioøkonomiske faktorer i et større perspektiv, kan slike forhold også knyttes til stabilitet og materiell trygghet. Respondentene verdsetter folk i arbeid og økonomisk trygghet i befolkningen. Dette har trolig sammenheng med at primærbehovene mat og et sted å bo kommer først. Relevansen av slike faktorer betyr at det i stor grad er innbyggernes levestandard som, i alle fall opp til primærbehovgrensen er nådd, vektlegges når statsministernes prestasjoner blir vurdert.

DW-verdien på 2.052 innebærer ingen autokorrelasjon. Det er altså ikke tendenser til at de variabelverdiene har sammenheng med tidligere verdier av seg selv.

Toleranseverdiene er godt over det kritiske 0.5-nivået. Dette betyr at kollinearitet ikke er et problem mellom variablene i modellen.

Av tabellen ser vi at prediksjonen av *BNP endring* på storhet er som forventet. En økning på ett prosentpoeng predikerer 0.23 poeng høyere storhetsgrad. Kontrollert for arbeidsledighet styrkes betaverdien. *BNP endring* har likevel høy p-verdi til at den foreløpig kan inkluderes i den endelige modellen.

Når gjennomsnittlig *arbeidsledighet* øker med ett prosentpoeng predikerer modellen en nedgang i storhetsgrad på -0.126 poeng. Variabelen er svært nær det kritiske signifikansnivået. Ut fra analysens definerte kriterier vil det bli vurdert om den skal testes i den endelige modellen. *Arbeidsledighet* er et antatt teoretisk viktig mål, p-verdien er svært nær signifikansnivået og den viste signifikant sammenheng med storhet i den bivariate analysen (0.084). Variabelens rolle i resten av analysen vil bli diskutert etter test i *backward regression*.

5.3.4 Biografiske effekter

Den bivariate analysen indikerte at ingen av de biografiske variablene har reell betydning for storhetsgraden til norske statsministere, men *utdanning over grunnskole* var likevel nær signifikansnivået. Modellen med biografiske variabler er vist i tabell 6.

Tabell 6 – biografiske faktorer

Variabel	b	St.feil til reg. koef.	beta	p- verdi	toleranse
Konstant	6.171	2.305		0.015	
Utdanning over grunnskole	-0.686	0.610	-0.266	0.276	0.824
Utdanning tilsvarende grunnskole	0.258	0.922	0.075	0.783	0.645
Region Sør-Vest Norge	0.351	0.801	0.111	0.666	0.718
Region Nord-Norge	0.197	0.820	0.057	0.813	0.816
Alder ved innsettelse	-0.047	0.041	-0.291	0.267	0.713
Leveår etter avsettelse	-0.008	0.029	-0.069	0.762	0.908

R²=0.171 justert R²= -0.106 F=0.618 F.sign=0.714 DW=1.645 St.feil til regresjonen=1.358

Modellen har en forklart varians på 17.1 prosent, mens justert R² er -0.106. Når det tas hensyn til antall prediksjonsvariabler og observasjoner, fører det altså til en kraftig reduksjon i modellens prediksjonskraft, og dette er et svakhetstegn. Standardfeilen til regresjonen er på 1.358. Verdien er vesentlig høyere enn i de to foregående modellene og innebærer at den biografiske modellen er den dårligste til å predikere storhet.

F-verdien er lav og ikke i nærheten av å være signifikant. Dette er ikke overraskende med tanke på at ingen av variablene er signifikante. DW-verdien på 1.645 er forholdsvis lav, men godt over det kritiske 1.5-nivået.

Samtlige variabler har toleranseverdier på over 0.6. Kollinearitet er altså ikke et problem i modellen. Oppsummert kan man si at modellen virker både svak og ubetydelig for prediksjon av storhet.

Standardfeilene til samtlige regresjonskoeffisienter er høyere eller like under b-verdien, noe som innebærer t-verdiene er forholdsvis langt under 1.71-grensen. Dette bekreftes av p-verdiene.

Når det gjelder type utdanning har *over grunnskole* negativ prediksjon mens *tilsvarende grunnskole* har positiv. En statsminister med litt utdanning blir altså vurdert lavere enn dem

med hovedfag. Regjeringsledere med kun grunnskole får likevel høyere score enn dem med hovedfag. Dette er motsatt av hypotesen om at mer utdanning har positiv prediksjon på storhetsvurderingen. Årsaken kan være en "Askeladden-effekt" der en statsminister som skiller seg ut fra mengden, og har et antatt dårligere utgangspunkt for å gjøre en god jobb enn dem med høyere utdanning, blir vurdert mer positivt. På tross av at *utdanning over grunnskole* ikke er signifikant, vil det være interessant å se resultatene fra testen *backward regression*.

Variablene *region Sør-Vest* og *region Nord-Norge* har positiv prediksjon, men er langt fra signifikante. Prediksjonen er overraskende i forhold til antakelsen om at jo lengre bort fra *Østlandet* man kommer, jo vanskeligere er det å påvirke det etablerte sentrum. I hypotesekapittelet ble det likevel knyttet usikkerhet til retningen på prediksjonen. Verdien kan innebære at det blir vurdert positivt å ha tilhørighet til periferien og utfordre det etablerte sentrum med dette utgangspunktet.

Alder ved innsettelse har forventet prediksjon. Det er trolig mer sannsynlighet at en statsminister som har høy alder ved innsettelse blir sittende en kortere periode med makten enn en yngre person. Det er sannsynlig at denne variabelen påvirker muligheten for å bli sittende lenge som statsminister og sette sitt preg på politikken. Lengde på statsministerperiode er som tidligere nevnt en av de antatt viktigste predikatorene på storhetsgrad. *Alder ved innsettelse* har lavest p-verdi i modellen, men er likevel et stykke over 0.10-grensen. Derfor vil den foreløpig ikke bli inkludert i den endelige modellen.

Leveår etter avsettelse har også negativ prediksjon. I hypotesekapittelet ble det antatt at prediksjonseffekten ville være positiv fordi man ved å leve lenge etter avsettelse har større mulighet til å påvirke sitt ettermæle gjennom for eksempel selvbiografier og deltakelse i den offentlige debatten. Hansson og Teigene bruker Einar Gerhardsen som eksempel på en statsminister som i stor grad har greid å påvirke sitt ettermæle etter han ble avsatt (1992: 323). Den høye p-verdien viser at variabelen likevel er langt fra signifikant.

Generelt er variablenes signifikansresultater, og den lave forklarte variansen, som forventet i forhold til resultatene fra den internasjonale litteraturen (Cohen 2004; Schlesinger jr. 1997; Simonton 1986: 262-267). Modellen bekrefter således at variablene trolig ikke har noen effekt på norske statsministers prestasjoner.

5.3.5 Samlede effekter

Antall måneder var den eneste variabelen som viste seg å være signifikant fra de foregående regresjonsanalysene. I tillegg er det flere variabler som er i nærheten av å være signifikante. Dette gjelder spesielt *BNP endring* og *arbeidsledighet*. Test på *backward regression* vil gi en indikasjon på hvilke variabler som bidrar signifikant til prediksjon av storhet.

Resultatet viser at variablene som ikke blir tatt ut av modellen i den stegvise elimineringsprosessen er *antall måneder*, *utdanning over grunnskole* og *arbeidsledighet*.

Arbeidsledighet var med p-verdien 0.117 nesten signifikant fra de gruppevise regresjonsanalysene, og den var signifikant i bivariat (0.084) og *backward* regresjon. Ifølge Gelman og Stern er ikke forskjellen mellom signifikante og insignifikante verdier nødvendigvis signifikant, og man skal derfor være åpen for å inkludere insignifikante variabler fra multivariat regresjon (2006: 328). *Arbeidsledighet* har også høy toleranseverdi. Ut fra kriteriene diskutert i forrige kapittel oppfyller variabelen således to av punktene, og er svært nær ved å oppfylle kriteriet om statistisk signifikans. På grunnlag av dette blir den analysert i den endelige modellen.

Det knyttes usikkerhet til den substansielle signifikansen av variabelen *utdanning over grunnskole*. Statsministerne i denne analysen blir predikert positivt i forhold til *hovedfag* hvis de har utdanning *grunnskole*, men ikke hvis de har *over grunnskole*, noe som virker ulogisk. I tillegg har *utdannelse* lav validitet på grunn av den sterke økningen i utdanningsnivået fra 1960-tallet. Utviklingen gjør at statsministere etter denne perioden har generelt høyere utdannelse enn de andre, og variabelverdiene har således mindre relevans for problemstillingen. I tillegg var variabelen insignifikant i både bivariat (0.113) og multivariat regresjon (0.276). Den blir derfor ekskludert fra videre analyse.

På grunnlag av diskusjonen over skal nå *antall måneder* og *arbeidsledighet* analyseres i den endelige modellen. Denne er vist i tabell 7. Ved å se på betaverdiene kan man sammenligne prediksjonseffekten til de to variablene. Ut fra de gruppevise analysene vil trolig *antall måneder* ha størst prediksjon på storhetsgraden til norske statsministere.

Tabell 7 – endelig modell

Variabel	b	St.feil til reg. koeff.	beta	p- verdi	toleranse
Konstant	2.871	0.330		0.000	
Antall måneder	0.017	0.004	0.674	0.000	0.978
Arbeidsledighet	-0.100	0.057	-0.251	0.090	0.979

R²=0.568 justert R²=0.529 F=14.457 F.sign=0.000 DW=1.856 St.feil til regresjonen:0.886

Tabellen viser en justert R² på 0.529. Faktorene predikerer med andre ord over halvparten av variasjonen i storhetsgraden blant norske statsministere. Standardfeilen til regresjonen er 0.886, og dette er lavest av samtlige modeller. Det at modellen predikerer storhet best er som forventet med tanke på at det er de mest relevante variablene som ble inkludert fra analysene av de tre variabelgruppene.

F-testen gir verdien 14.457. Det innebærer at den endelige modellens samlede effekt er vesentlig større enn i de foregående. I forhold til variablene den inneholder er dette som forventet. F-testen er signifikant, og det betyr at modellen som helhet viser en lineær sammenheng ulik fra null, altså at den har effekt på storhetsgrad. Autokorrelasjon er ikke et problem i den endelige modellen. Da den kun består av to variabler, og korrelasjonsmatrisen allerede er diskutert, skal ikke kollinearitet være et problem. Toleranseverdiene bekrefter dette.

Konstantleddet har b-verdi på 2.871. Ser vi på de to variablene separat er b-verdien til *antall måneder* 0.017 mens betaverdien er 0.674. Modellen predikerer at en måned lenger som statsminister gir en økning i rangeringen på 0.017 score. Ser man på prediksjonseffekten av å sitte ett år lenger som statsminister, $0.017 \cdot 12$, er denne 0.2 poeng. Etter fem år vil regjeringslederen få 1 storhetsscore høyere enn en som var innsatt i én måned. Variabelen er svært signifikant, og man kan med nær 100 prosent sikkerhet forkaste nullhypotesen som sier at *antall måneder* ikke predikerer storhetsgrad. Prediksjonseffekten av hvor lenge statsministeren er innsatt samsvarer i stor grad med funn fra ulike studier av amerikanske presidenters prestasjoner (Simonton 1986: 270). På tross av at disse har inkludert andre uavhengige variabler i sine modeller, og at storhetsgrad ofte er operasjonalisert litt annerledes enn i denne analysen, er likheten interessant. I alle fire aktuelle studier fra USA har *antall år som president* en signifikant prediksjon på storhet. Dette styrker antakelsen om at variabelen har effekt på tvers av stater og politiske systemer.

Respondenten Finn Olstad argumenterer på følgende måte for sin vurdering av Gerhardsen som en stor statsminister: ”Han kan sies å stå i en klasse for seg i vårt århundre, både ut fra sin personlighet, sin politiske begavelse og lange funksjonstid” (Cordtsen 1997a). Dette indikerer også at respondentene vektlegger denne variabelen i sine vurderinger.

Gjennomsnittlig *arbeidsledighet* har b-verdi på -0.100 . Dette innebærer at en statsminister som regjerte under perioder med åtte prosent arbeidsledighet får 0.5 lavere storhetsscore enn en som regjerte med tre prosent ledighet. Variabelens p-verdi er 0.090, den er altså signifikant på ti prosent nivå med to-halet test. Verdien innebærer at man kan forkaste nullhypotesen med 90 prosent sikkerhet.

En sammenligning av variabelenes betaverdi viser at *antall måneder* som statsminister, 0.674, har mer enn dobbelt så stor prediksjonseffekt som arbeidsledighet, -0.251 , på storhetsgrad. I forhold til den internasjonale litteraturen er det som forventet at lengden en statsminister regjerer er den viktigste faktoren for å predikere deres prestasjoner (Murray & Blessing 1994: 67).

I et større perspektiv kan betydningen av regjeringstid ses i sammenheng med politisk stabilitet, samt respekt og tillit i befolkningen. Innbyggerne påvirker gjennom valg hvor lenge personen blir sittende³⁴. Dermed kan man anta at en statsminister som over lengre perioder har hatt støtte blant velgerne, blir vurdert bedre enn dem som ble sittende i korte perioder. Når respondentene også vektlegger denne faktoren, er det en indikasjon på at det er samsvar mellom deres og befolkningens vurderinger av hvilke kriterier som bør ligge til grunn for å bli en stor statsminister. Dette støttes av Jeffrey E. Cohen selv om han fokuserer på *informed citizens*³⁵ fremfor hele befolkningen. Cohen mener at [e]xperts and informed citizens rate presidents similarly and use similar criteria” (2004: 913). Betydningen av *antall måneder* i denne analysen er verdifull i forhold til forskningsfeltet, fordi det styrker antakelsen om at storhetsscoren også gjenspeiler velgernes oppfatning om hvilke kriterier som bør ligge til grunn for vurderingene.

³⁴ Det kan diskuteres hvor stor betydning befolkningen har på regjering og hvem som blir statsminister fordi det ikke nødvendigvis er de partiene med størst oppslutning som danner grunnlag for regjering (Østerud *et al* 2003: 87). Likevel må det kunne sies at velgerne har en vesentlig effekt på statsminister og regjering.

³⁵ Undersøkelsen tok utgangspunkt i C-span, en amerikansk kringkastingsstasjon som er viet til kontinuerlig dekning av offentlige myndigheter og begivenheter. C-span serne, som Cohen mener er mer opplyste enn den generelle befolkningen, fikk mulighet til å vurdere presidentenes prestasjoner på de samme dimensjonene som ekspertene.

Man kan også se sammenhengen mellom *antall måneder* og storhet i forhold til at flere måneder som statsminister øker sjansen for at historisk viktige hendelser inntreffer. I den amerikanske litteraturen blir sistnevnte regnet for å være en forutsetning for å bli betegnet som stor (Simonton 1986: 279). Theodore Roosevelt mente i denne sammenheng at han ikke fikk muligheten til vise sin styrke fordi den historiske konteksten aldri ga ham sjansen (sitert i Adler 2003: 6). Dette samsvarer med Crocketts argument om at en president er avhengig av viktige hendelser for å bli betegnet som stor, og at tilfeldigheter derfor blir et viktig element i prestasjonsvurderingene (Crockett 2002: 240-242).

Overensstemmelsen mellom storhetsstudier fra USA og Norge styrker verdien og troverdigheten av at denne analysen faktisk predikerer statsministerprestasjoner, og at det i mindre grad er respondentenes subjektive oppfatninger som påvirker resultatet.

Det at sosioøkonomiske forhold har betydning for storhet indikerer at stabilitet er en viktig forutsetning for å bli betegnet som stor. Høy arbeidsledighet er et tegn på at livskvaliteten til befolkningen er lavere enn når flere er sysselsatt. Dette tas i betraktning når rangeringene foretas. Det er en vanlig antakelse at sosioøkonomiske forhold i stor grad avhenger av internasjonale konjunkturer, og det kan derfor diskuteres hvor mye makt en statsminister har til å påvirke disse forholdene. Det er derfor interessant at respondentene likevel tillegger dette ansvaret til statsministeren.

5.4 Stabilitet – analyser av residualer

Med utgangspunkt i den endelige modellens resultater vil det være interessant å se om storhetsscoren til de ulike statsministerne samsvarer med hvor lenge de regjerte og hvor høy den gjennomsnittlige arbeidsledigheten var i deres periode. Tabell 8 indikerer hvor godt de ulike statsministerne blir predikert av den endelige modellen. Hvis en regjeringsleder for eksempel fikk lav score på tross av at personen satt lenge som statsminister, og det samtidig var lav arbeidsledighet, øker sjansen for at statsministerens prestasjon blir dårlig predikert av modellen.

Tabell 8 – Storhetsscore og variabelverdier

Statsminister	Score	Antall måneder	Arbeidsledighet
Christian Michelsen	5.75	32	1.1
Jørgen Løvland	2.63	5	1.2
Gunnar Knudsen	5.25	114	1.1
Wollert Konow	1.84	25	0.9
Jens Bratlie	2.18	11	0.6
Otto B. Halvorsen	2.78	16	1.7
Otto Blehr	2.58	21	6.6
Abraham Berge	2.17	14	4.5
Johan L. Mowinckel	3.85	86	8.3
Ivar Lykke	2.21	23	8.1
Christopher Hornsrud	3.56	1	7.3
Peder Kolstad	1.80	11	10.4
Jens Hundseid	1.45	12	9.8
Johan Nygaardsvold	4.50	124	7.1
Einar Gerhardsen	5.85	209	0.9
Oscar Torp	3.00	39	0.9
John Lyng	4,18	1	1.3
Per Borten	3.50	66	0.9
Trygve Bratteli	3.85	48	0.9
Lars Korvald	3.16	12	0.8
Odvar Nordli	2.9	61	1.1
Gro H. Brundtland	4.95	123	3.0
Kåre Willoch	4.4	56	2.6
Jan P. Syse	2.21	13	4.3
Thorbjørn Jagland	1.68	12	4.0

Når man tolker tabellen over er det viktig å ta i betraktning at lengden en statsminister var innsatt har over dobbelt så stor effekt som arbeidsledighet i den endelige modellen. Oversikten viser at Christian Michelsens storhetsscore er veldig høy tatt i betraktning at han var innsatt relativt kort tid, og man kan derfor anta at han blir dårlig predikert av modellen. Hvis man sammenligner Michelsen med Wollert Konow ser man at de har relativt like verdier på de to variablene, og likevel er storhetsscoren på henholdsvis 5.75 og 1.84. Dette

betyr at Konow trolig også faller utenfor modellen fordi det var mange som var innsatt kortere enn ham, regjerte med høyere arbeidsledighet og samtidig blir vurdert høyere. Odvar Nordli er en annen statsminister som har fått lavere score enn modellen skulle tilsi er. Han regjerte i 61 måneder, og det var gjennomsnittlig kun 1.1 prosent arbeidsledighet i perioden. Likevel får han ikke høyere score enn 2.9. John Lyng blir trolig dårlig predikert av modellen. På tross av at han kun var innsatt i én måned har han en storhetsscore på 4.18. Selv om ledigheten i perioden var lav burde scoren vært lavere hvis respondentene hadde vektlagt innsettelsesperioden hans på samme måte som modellen.

En statsminister som sannsynligvis blir godt predikert av modellen er Otto Blehr fordi verdiene på de uavhengige variablene ser ut til å korrespondere med storhetsscoren respondentene har gitt ham. Han var innsatt i en forholdsvis kort periode og arbeidsledigheten var i tillegg relativt høy. Intuitivt virker derfor hans lave score å samsvare med variablenes effekter. Peder Kolstads prestasjon blir trolig også godt predikert av den endelige modellen. Han har blant de laveste storhetsscorene. Tatt i betraktning hans elleve måneder med makten, og høy arbeidsledighet i perioden, virker det logisk.

Med utgangspunkt i den endelige modellen skal det nå analyseres hvor godt de ulike statsministernes prestasjoner blir predikert. Residualverdiene blir vist i tabell 9. Da målet med denne oppgaven er å predikere mest mulig av den avhengige variabelen, vil det i residualanalysen diskuteres variabler som kan tenkes å øke prediksjonen på de aktuelle statsministernes storhetsgrad.

Tabell 9 – Hvem faller utenfor modellen?

Statsminister	Standardisert residual
Christian Michelsen	2.75
John Lyng	1.60
Christopher Hornsrud	1.58
Wollert Konow	-1.54
Thorbjørn Jagland	-1.12
Odvar Nordli	-1.02
Kåre Willoch	0.94
Jens Bratlie	-0.92
Jens Hundseid	-0.73
Gunnar Knudsen	0.61
Einar Gerhardsen	-0.57
Abraham Berge	-0.55
Oscar Torp	-0.51
Jan P. Syse	-0.51
Per Borten	-0.46
Johan L. Mowinckel	0.38
Gro H. Brundtland	0.31
Trygve Bratteli	0.28
Ivar Lykke	-0.27
Peder Kolstad	-0.24
Johan Nygaardsvold	0.24
Jørgen Løvland	-0.23
Otto B. Halvorsen	-0.22
Lars Korvald	0.18
Otto Blehr	0.01

Tabellen viser statsministernes residualverdier i kronologisk rekkefølge fra størst til minst. Hvis flere av de avvikende observasjonene viser fellestrekk kan "[e]n eller flere forklaringsvariabler muligens fange opp disse trekkene" (Midtbø 2007: 89). De som har høyere/lavere verdier enn det første kuttunktet $1/-1$ vil bli studert nærmere. Det andre kuttunktet går på $2/-2$, og blir brukt fordi det er større sjanse for at disse statsministrene avviker signifikant fra de andre observasjonene. For å begrense omfanget av analysen vil det kun inkluderes nye konkrete prediksjonsvariabler på statsministere med residualverdier over/under $2/-2$.

Seks statsministere skiller seg ut i tabellen. Christian Michelsen er som forventet underpredikert i den forstand at hans faktiske verdi er høyere enn den predikerte ($e=Y_f-Y_p$). Med residualverdien 2.75 er han den som blir dårligst predikert av den sosioøkonomiske situasjonen og hvor lenge han var innsatt. Det betyr at den høye scoren hans i stor grad blir predikert av variabler som ikke er med i modellen og innebærer at modellen estimerer en høyere storhetsscore enn det den faktiske verdien tilsier. Ut fra betydningen av de to aktuelle

variablene burde han med andre ord fått lavere score enn det respondentene har gitt ham. John Lyng blir underpredikert med 1.60 og Christopher Hornsrud med 1.58. Wollert Konow er med verdien -1.54 den neste statsministeren som har høy residualverdi, men han blir i motsetning til de to sistnevnte overpredikert. Thorbjørn Jagland og Odvar Nordli blir også overpredikert, henholdsvis med verdiene -1.12 og -1.02 .

Det ble i teorikapittelet forventet at Einar Gerhardsen ville bli underpredikert i den endelige modellen som følge av å være "landsfader" og stå for oppbygningen av landet etter Andre verdenskrig. Tatt i betraktning at han regjerte svært lenge, og at det i tillegg var svært lav arbeidsledighet i etterkrigsårene, er det likevel ikke overraskende at hans prestasjon blir relativt godt predikert i den endelige modellen.

5.4.1 Hvorfor avviker noen statsministere?

Årsaken til at Christian Michelsen blir dårligst predikert av modellen har trolig sammenheng med statusen han fikk etter løsrivelsen fra Sverige i 1905. I Simonton-studien av amerikanske presidents storhet ble George Washington predikert høyere av historikerne enn modellen skulle tilsi (1986: 272). Som nevnt i teorikapittelet kan det trekkes paralleller mellom Washington og Christian Michelsen. Begge satt ved makten da de respektive statene ble selvstendige, og de blir regnet for å være henholdsvis første president og statsminister.

Michelsen er den eneste statsministeren med residualverdier som overgår det andre kuttpunktet på $2/-2$. Ved å legge inn en dummyvariabel på Michelsen kan man se hvor mye han som person påvirker modellen. Det viser seg at dummyen har p-verdi 0.002. Han er altså signifikant forskjellig fra de andre statsministrene i analysen. B-verdien er 2.627 mens betaverdien er 0.407. Resultatene er i samsvar med residualanalysen, og innebærer at Michelsens effekt på storhet i den endelige modellen er 2.63 score. Dette styrker antakelsen om at unionsoppløsningen i 1905 har en effekt på storhetsscoren hans. I tillegg argumenterer mange³⁶ av respondentene i Cordtsen-undersøkelsen for omstendighetene rundt 1905 som årsak til at de gir ham høy score. Michelsen var i tillegg innsatt for en relativt kort periode, og dette er også noe av årsaken til at modellen predikerer ham dårligst. Hvis Michelsen hadde blitt vurdert ut fra modellen ville han fått scoren 3.12 og blitt regnet for å være en middelmådig statsminister.

³⁶ Blant annet Nils Johan Ringdal, Rolf Danielsen, Hans Fredrik Dahl, Olav Versto, Finn Olstad og Bernt Hagtvet vektlegger dette.

Tester av de fire andre statsministerne som nådde det første kuttunktet viste at de ikke var signifikant forskjellige fra resten av observasjonene. Det at Michelsens prestasjon i stor grad blir bedømt ut fra unionsoppløsningen, gjenspeiler en svakhet ved respondentene fordi det innebærer at de ikke opererer med generelle kriterier i statsministervurderingene. Ved for eksempel å be respondentene vurdere prestasjon etter gitte kriterier, vil man unngå dette i fremtidige analyser.

I forhold til forutsetningen om normalfordelte restleddverdier er Michelsen en uteligger. På tross av dette er ikke observasjonens verdier et direkte brudd på forutsetningen fordi normalfordelingen ellers ikke viser store avvik (for test av forutsetningene se appendiks E). Ved å legge inn dummy for Michelsen som person er det testet og funnet effekt av observasjonen, og således er problemet løst statistisk. Likevel vil ikke dummyvariabelen for Michelsen inkluderes i den endelige modellen. Det er modellens variabler som er substansielt viktige i forhold til analysens formål, og dette overgår verdien av statistisk ”korrekte” og maksimerte resultater.

Av tendenser i residualene ser man av tabell 6 at Christopher Hornsrud og John Lyng blir underpredikert med omtrent samme verdi i den endelige modellen. Dette har sammenheng med at begge statsministerne satt med makten én måned hver. Tatt i betraktning at lengden regjeringslederne var innsatt er den viktigste predikatoren på storhet, passer de dårlig inn i modellen. Respondenten Olav Verstos argument for å rangere Lyng i kategorien *over middels* er betegnende for det mange av respondentene la vekt på: ”Han regisserte forspillet til ettpartistatens fall med sin korte visitt ved makten i 1963 og fikk vist at de borgerlige ville og kunne regjere sammen” (Cordtsen 1997a). Dette vektlegges og i Cordtsens artikkel (1997b: 24).

Ut fra respondentenes kommentarer kan det virke som at Christopher Hornsrud blir rangert såpass høyt, 3.56, fordi han representerte et veiskille i norsk politikk i 1927. Ifølge Bernt Hagtvet gjorde Hornsrud det ”[m]ulig å tenke seg en sosialistregjering i Norge” (Cordtsen 1997a). Nils Johan Ringdal støtter også opp om dette argumentet. Det foregående kan knyttes til variabelen *partipolitisk skillelinje*, som var inspirert fra Rokkan, selv om denne viste seg å være insignifikant i analysen (1966: 85-89).

Det er forøvrig en tendens til at statsministerne som blir dårlig predikert av modellen var innsatt i relativt korte perioder. Mens gjennomsnittet satt 45 måneder med makten var kun en

av de aktuelle statsministre, Odvar Nordli, innsatt lenger enn 25 måneder. Det kan likevel ikke trekkes likhetstrekk mellom alle regjeringslederne fordi for eksempel Thorbjørn Jagland satt kort tid og fikk lav score, noe som er logisk ut fra den endelige modellen. Det er vanskelig å se klare sammenhenger mellom statsministere som blir dårlig predikert i modellen. De er fra ulike partier, spredt over ulike perioder av 1900-tallet og modellen fanger allerede opp arbeidsledighet og *antall måneder*. Dette gjør det vanskelig å knytte kontekstuelle forhold til de høye residualverdiene.

Det knyttes usikkerhet til hvorfor Wollert Konow ikke får høyere score av respondentene. Olav Versto vurderte Konow i kategorien *under middels* fordi regjeringen hans var svak. Dessuten var det dårlig samarbeid mellom Frisinnede Venstre og Høyre. Både Versto og Nils Johan Ringdal nevner hans provoserende uttalelser om sidemålsstilen og målsaken som en annen årsak til at de bedømmer han lavt (Cordtsen 1997a).

Respondentene vurderer Thorbjørn Jagland lavere enn modellen predikerer. Dette kan ha sammenheng med at han er den nyeste statsministeren i utvalget. I teorikapittelet ble det fokusert på at nyere presidenter i USA generelt blir dømt hardt. Hendelser og beslutninger virker mer dramatiske i det de skjer enn når respondentene ser statsministre i et større og mer nyansert perspektiv. I ettertid får de ofte mer informasjon, gjennom blant annet offentlige dokumenter som blir frigitt, om bakgrunnen for beslutninger som ble tatt (Lønnstrøm & Kelly 2003: 631). Eisenhower er et eksempel på dette i USA. Han har gradvis steget fra 27. plass i Schlesingers 1962-undersøkelse til niende plass i to ulike studier i henholdsvis 1982 og 1996. Årsaken til dette var at respondentene fikk mer informasjon om konteksten han regjerte under, og bakgrunnen for valgene han foretok (Schlesinger jr. 1997: 182). Respondenten Tore Pryser uttaler i forbindelse med vurderingen av Jagland at hans innsats i dag fortoner seg som en fiasko, men at "[v]i først etter en tid får et tydelig bilde av konsekvensene av vedkommendes politikk" (Cordtsen 1997a).

Flere respondenter nevner økonomiske rot og internasjonal nedkonjunktur som årsak til hvorfor de vurderer Odvar Nordli som middelmådig. Finn Olstad er en av dem som mener at Nordlis forsøk på motkonjunkturpolitikk ikke kunne fungere i lengden. "Hans forsøk på å føre videre tradisjonell sosialdemokratisk politikk under nye forhold måtte ende i katastrofe" (Cordtsen 1997a). Noen respondenter beskriver ham også som en anonym leder uten ambisjoner. Ifølge Hans Fredrik Dahl manglet Nordli egenskaper til å virke samlende på

regjeringen. ”En mann uten egenskaper annet enn det absolutte middels” (Cordtsen 1997a). Handlingskraft er som tidligere nevnt regnet for å være en viktig faktor i vurderingene av amerikanske presidenter. De nevnte forholdene kan være noe av årsaken til at Nordli blir vurdert lavere av respondentene enn det den endelige modellen predikerer.

Når det gjelder innflytelsesrike case så vil det nå fokuseres på $sdf\beta$ og $sdf\hat{t}$ -verdier (for fullstendig oversikt se appendiks A). Variablene som er med i den endelige modellen er *antall måneder og arbeidsledighet*. $sdf\beta$ -verdier på *antall måneder* viser at Einar Gerhardsen med -0.692 er den eneste som overstiger det kritiske 0.4 -nivået nevnt i kapittel 4. Med tanke på at han var innsatt i 205 måneder mens gjennomsnittet satt 45 måneder, er ikke dette overraskende. Resultatet betyr at Gerhardsen påvirker estimatet uforholdsmessig mye målt gjennom t -verdien som minker fra 4.557 til 3.865 , eller med 15 prosent, når statsministeren ekskluderes. Ingen av statsministerne hadde $sdf\beta$ -verdier over/under $0.4/-0.4$ på *arbeidsledighet*.

Den kritiske $sdf\hat{t}$ -verdien kan nå regnes ut fordi vi vet at det er to variabler i den endelige modellen. Gjennom formelen $2*\sqrt{p/N}$ der p er antall uavhengige variabler og N er antall observasjoner får vi kritisk verdi på $0.69/-0.69$. Statsministere som får verdier over/under $0.69/-0.69$ vil bli analysert nærmere. Observasjonenes $sdf\hat{t}$ -verdier viser at Christian Michelsen med verdien 0.97 overstiger det kritiske 0.69 -nivået. Han påvirker dermed modellens samlede effekt uforholdsmessig mye. Hvis Michelsen hadde blitt ekskludert fra modellen ville F -verdien økt fra 13.416 til 14.451 og forsterket sannsynligheten for at nullhypotesen, om at det ikke er en lineær sammenheng i populasjonen, er falsk. Michelsen kan således beskrives som en innflytelsesrik observasjon. Dette har sammenheng med at han ble rangert svært høyt på tross av at han var innsatt kort tid som statsminister. Einar Gerhardsen får $sdf\hat{t}$ -verdi på -0.75 og han kan dermed også sies å påvirke den endelige modellens samlede effekt uforholdsmessig mye, om enn i mindre grad enn Michelsen. F -verdien ville minket fra 13.416 til 12.276 ved å ekskludere Gerhardsen. Årsaken til dette ligger i hans eksepsjonelt lange regjeringstid som fører til at han som person påvirker modellens samlede effekt.

Diskusjonen om avvikende og innflytelsesrike statsministere avdekker ingen alvorlige svikt i modellen. Analysen identifiserer ingen periodeeffekter eller åpenbare utelatte variabler som skulle tilsi endringer i den enkle prediksjonsmodellen med to variabler. Det konstanteres at

Michelsen blir kraftig underpredikert, men en prediksjonsmodell kan ikke være case-basert og derfor blir dette registrert uten at det gir grunnlag for endringer i modellen.

5.5 Prediksjon av nye statsministere

Av statsministere etter 1905 er Kjell Magne Bondevik og Jens Stoltenberg de to som ikke er inkludert i denne analysen. Årsaken er at Cordtsen foretok sin studie i 1997. Vi kan nå anvende den endelige modellen til å predikere deres storhetsgrad.

Statsministernes verdier på variabelen *arbeidsledighet* er hentet fra Historisk statistikk 2007³⁷. For Bondevik blir den gjennomsnittlige arbeidsledigheten over seks år 4.0 prosent. For Stoltenberg blir gjennomsnittlig arbeidsledighet, definert ut fra kriteriene i metodekapittelet, regnet ut på grunnlag av årene han var innsatt frem til og med 2007. Han får da et gjennomsnitt på 3.3 prosent.

Når det gjelder *antall måneder* var Bondevik statsminister i 79 måneder. Stoltenbergs verdi på denne variabelen forutsetter at han blir sittende ut denne stortingsperioden. Han vil da ha vært statsminister i 69 måneder³⁸. Den endelige prediksjonsmodellen i analysen er definert som: $\text{prediksjon} = a + (b1 * v1) + (b2 * v2)$ der a er konstantleddet, $b1$ og $b2$ står for ustandardisert regresjonskoeffisient av henholdsvis *antall måneder* og *arbeidsledighet*, $v1$ står for antall måneder som statsminister og $v2$ står for gjennomsnittlig arbeidsledighet. Resultatene viser at Kjell Magne Bondevik vil få scoren 3.81 mens Jens Stoltenberg vil få 3.71. Hvis modellen hadde gjenspeilet virkeligheten ville Bondevik og Stoltenberg blitt rangert på henholdsvis tiende og ellefte plass³⁹ over Norges største statsministere.

5.6 Konklusjon av empiriske funn

Denne analysen består av to hoveddeler. Målet med den første delen var å predikere mest mulig av storhetsgraden til norske statsministere gjennom regresjonsanalyser. Kriteriene for hvilke variabler fra de institusjonelle, sosioøkonomiske og biografiske gruppene som skulle inkluderes i den endelige modellen var at de hadde substansiell signifikans, statistisk signifikans og høy toleranseverdi. Den andre delen av analysen hadde som formål å se hvilke statsministere som ble dårlig predikert av den endelige modellen. En diskusjon av årsakene

³⁷ Variabelverdiene er hentet fra Samfunnsveven (<http://www.nav.no/page?id=1073743220>).

³⁸ Se <http://samfunnsveven.no/TT/U40.cfm>.

³⁹ For kronologisk rangering se appendiks D.

til dette sto her sentralt. For statsministere med residualverdier over/under 2/-2 ville det inkludere nye variabler. Det skulle også analyseres hvilke statsministere som kunne betegnes som innflytelsesrike på resultatet, samt årsakene til dette. De nyeste statsministere, som ikke var en del av den opprinnelige analysen, ville deretter bli predikert ut fra den endelige modellen.

Den bivariate regresjonsanalysen indikerte at fire variabler hadde statistisk signifikant prediksjon på storhetsgrad. Disse var *antall måneder, flere statsministerperioder, gjenvalg og arbeidsledighet*. På grunn av at førstnevnte var den eneste av de institusjonelle variablene som var signifikant, kunne denne inkluderes i den endelige modellen på tross av lav toleranseverdi. Av de sosioøkonomiske variablene hadde *arbeidsledighet* p-verdi 11.7. Denne ble sett på som en substansielt viktig prediksjonsfaktor. I tillegg hadde variabelen signifikant effekt i bivariat og *backward* regresjon. Derfor ble den inkludert i den endelige modellen. De biografiske faktorene hadde liten betydning for storhetsgraden, men dette var ikke overraskende når man ser på funn fra den internasjonale litteraturen. Som Simonton sier det: “It is curious how little of presidential greatness can be rooted, whether directly or indirectly, in the president’s personal attributes” (Simonton 1986: 275).

Med justert forklart varians på 52.9 prosent kan den endelige modellen sies å ha vesentlig prediksjonskraft på storhetsgraden til norske statsministere. Med tanke på at forskningsfeltet er relativt nytt i Norge, og at analysen i denne sammenheng er unik, må modellen sies å ha større prediksjonskraft enn forventet. Standardfeilen til regresjonen i den endelige modellen var lavere enn i noen av de andre, noe som styrker funnene. Hvor lenge en statsminister var innsatt er den klart viktigste predikatoren på storhet. Ved å sitte ett år lenger med makten oppnår statsministerne en økning i storhet på 0.2 poeng. Med tanke på den store variasjonen i statsministernes regjeringstid, gir variabelen store utslag på storhetsgraden. Det er også interessant at effekten av denne faktoren er svært lik funnene i den amerikanske litteraturen der slike type studier har lang tradisjon. Dette styrker troverdigheten til at variabelens prediksjonseffekt er reell og at politisk stabilitet, samt respekt og tillit i befolkningen har betydning for respondentenes vurderinger. På grunn av at statsministerens regjeringstid avhenger av velgernes tillit er det sannsynlig at deres oppfatninger om storhet samsvarer med respondentenes vurderinger. Dette øker sjansen for at respondentenes subjektive oppfatninger ikke påvirker resultatet. I tillegg har makroøkonomiske forhold målt gjennom

arbeidsledighet signifikant prediksjon på storhet. Dette betyr at innbyggernes levestandard er sentral når respondentene vurderer statsministerprestasjoner.

Christian Michelsen, John Lyng, Christopher Hornsrud, Wollert Konow, Thorbjørn Jagland og Odvar Nordli var de seks statsministere som ble dårligst predikert av modellen. Førstnevnte ble vurdert høyere av respondentene enn den endelige modellen tilsa. Dette var som forventet fordi Michelsen i kapittel 2 ble knyttet opp mot unionsoppløsningen i 1905 og status som landsfader. Residualanalysen viste også at personen Michelsen er signifikant forskjellig fra de andre statsministerne. Respondentenes kommentarer tyder på at det i stor grad er løsrivelsen fra Sverige som ligger til grunn for den høye scoren. På tross av John Lyngs svært korte regjeringstid får han relativt høy score. Respondentene argumenter for at det er hans rolle i å få slutt på Arbeiderpartiets ettpartistat som er årsaken til dette. Det knyttes mer usikkerhet til hvorfor Christopher Hornsrud og Wollert Konow blir dårlig predikert, men ut fra respondentenes kommentarer det kan virke som sistnevnte blir straffet for å ha vært en svak leder. Hornsrud får trolig sin relativt høye score fordi han representerte et veiskille i norsk politikk i forhold til at tanken om en fungerende sosialistisk regjering ble virkeliggjort. I den internasjonale litteraturen er det funnet tendenser til at nyere presidenter blir vurdert strengere enn dem som satt lengre tilbake i tid. Noe av årsaken til Thorbjørn Jaglands lave score kan trolig forklares av dette. Det kan virke som Odvar Nordli blir vurdert lavt på grunn av internasjonale nedkonjunkturer og den måten han taklet dette på. I tillegg mener flere respondenter at han var lite handlingskraftig og ikke hadde egenskapene som trengtes for å få til en slagkraftig regjering. I amerikanske studier blir handlingskraft betegnet som en viktig egenskap for å bli en stor president (Maranell 1970). Noe overraskende hadde ikke Einar Gerhardsen høye residualverdier. Det ble i kapittel 2 antatt at han ville bli kraftig underpedikert som følge av status som landsfader i forbindelse med gjenoppbyggingen av Norge og opprettelsen av velferdsstaten etter 1945.

Når det gjelder innflytelsesrike case ble det gjennom sdffit-verdiene bekreftet at Christian Michelsen og Einar Gerhardsen påvirker modellens samlede effekt uforholdsmessig mye. Førstnevnte på grunn av unionsoppløsningen og sistnevnte på grunn av sin lange regjeringstid. Gerhardsens store effekt på *antall måneder* er på grunn av hans overlegne 209 måneder som statsminister.

Prediksjonen av de nye statsministerne Kjell Magne Bondevik og Jens Stoltenberg ga storhetsgrad på henholdsvis 3.81 og 3.71.

6. KONKLUSJON

Det overordnede forskningsspørsmålet i denne studien er: *Hvilke faktorer predikerer storhetsgraden blant norske statsministere i perioden 1905-1997?* Analysen er inspirert av det store forskningsfeltet som presidentprestasjoner i USA har utviklet seg til å bli siden oppstarten i 1948. På grunn av mangelen på lignende studier i Norge har fremgangsmåten i denne analysen vært å fokusere på en bred og grunnleggende tilnærming til temaet slik forskere også har gått frem i USA. I denne sammenheng var det naturlig å ta utgangspunkt i Cordtsen-studien. Arbeidet hans la forholdene til rette for at denne analysen kunne bli gjennomført innenfor rammene av en masteroppgave.

Prediksjonskapittelet tok først for seg den internasjonale litteraturen. Det ble her argumentert for at det er stort samsvar mellom resultater fra ulike amerikanske presidentstudier, spesielt blant personene i toppen og bunnen av kåringene. Dette styrker forskningsfeltet fordi det viser at vurderingene er mindre avhengig av respondentene som deltar. Det ble deretter fokusert på flere utfordringer knyttet til bruken av variabler som er tilpasset den amerikanske konteksten. Sartoris begreper *conceptual travelling*- og *stretching* sto her sentralt. Betydningen av tid og historisk kontekst er diskutert i den internasjonale litteraturen. Det ble først fokusert på årsaker til hvorfor en regjeringsleders storhetsscore endrer seg over tid, og deretter på hvorvidt man egentlig kan måle storhet når de kontekstuelle forutsetningene er ulike for hver statsministers regjeringstid. Theakston og Gills argument oppsummerer problematikken: "The problem with comparative scoring exercises is that no two leaders are ever dealt the same hand" (2006: 211). Det ble konkludert med var at respondentene trolig tar hensyn til de ulike historiske kontekstene statsministerne regjerer under, en antakelse som ble styrket i gjennomgangen av Cordtsen-respondentenes utfyllende kommentarer.

I hypotesekapittelet ble de inkluderte variablene presentert og diskutert. Disse dannet grunnlaget for hvilke faktorer som predikerer storhetsgraden til norske statsministere i perioden 1905-1997. Sentrale diskusjoner i dette kapittelet var inspirasjonen for variablene, presentasjon av hypoteser og antatt prediksjon på storhetsgrad. Det ble antatt at gruppen av institusjonelle variabler var mest relevante mens den biografiske gruppen ble vurdert som mindre relevant. *Leder av flertallsregjering (+)*, *antall måneder som statsminister (+)*, *flere statsministerperioder (+)* *gjenvalg (+)*, *BNP endring (+)* og *arbeidsledighet (-)* var enkeltvariablene ble det forventet størst prediksjonseffekt av.

Kapittelet *data, operasjonalisering og metode* tok først for seg datagrunnlaget i Cordtsen-studien. På grunn av at verdiene på avhengig variabel i analysen ble definert med utgangspunkt i denne studien var validiteten på dataene et sentralt tema. Det ble også diskutert mulige problemer knyttet til respondentenes objektivitet, et tema som har vært gjenstand for mye diskusjon i den internasjonale litteraturen. Deretter ble det fokusert på operasjonalisering av variablene som ligger til grunn for denne studien. Metodiske temaer i forhold til den kommende analysen sto sentralt i den tredje hoveddelen. Valgene om å foreta en kvantitativ analyse med multivariat regresjon, samt fokuset på prediksjon fremfor kausalitet, ble også forsvart i denne delen. Dessuten ble kriterier for variabelutvelgelse i analysen definert.

6.1 Oppsummering av funn

I den internasjonale litteraturen går *antall år/måneder som president/statsminister* igjen som en viktig prediksjonsfaktor på storhetsgrad (Simonton 1986; Theakston & Gill 2006: 205). I tillegg har *gjenvalg* stått sentralt i prediksjon av amerikanske presidenters prestasjoner (Simonton 1986: 268 og 273). Den bivariate analysen indikerte at *antall måneder (+)*, *flere statsministerperioder (+)*, *gjenvalg (+)* og *arbeidsledighet (-)* var sentrale for prediksjon av storhetsgraden til norske statsministere.

Det ble deretter foretatt multivariate regresjonsanalyser av de *institusjonelle*, *sosioøkonomiske* og *biografiske* variabelgruppene. *Antall måneder* var den eneste av *institusjonelle* variablene som hadde signifikant p-verdi. På tross av at variabelens lave toleranseverdi indikerte multikollinearitet ble ikke dette ansett som et problem da ingen andre uavhengige variabler var signifikante. Av de sosioøkonomiske variablene var verken *BNP endring* eller *arbeidsledighet* signifikante, men p-verdiene på henholdsvis 0.191 og 0.117 var nær 0.10-nivået. Begge ble ansett som substansielt signifikante faktorer, men hvorvidt de skulle inkluderes ville bli avgjort etter resultatene fra *backward regression*. Innenfor gruppen av biografiske variabler var det ikke overraskende ingen signifikante variabler. Dette er i samsvar med funn fra den internasjonale litteraturen (Cohen 2004; Schlesinger jr. 1997; Simonton 1986: 262-267).

Med *antall måneder* som eneste signifikante variabel ble det nå foretatt en ekstra statistisk sjekk av samtlige variabler i *backward elimination*. Denne elimineringsmetoden viste at også *arbeidsledighet* og *utdanning over grunnskole* bidro signifikant til prediksjon av storhet.

Sistnevnte variabel ble deretter ekskludert på grunn av lav substansiell signifikans i form av lav validitet. I tillegg var ikke variabelen signifikant i bivariat og multivariat regresjon. *Arbeidsledighet* ble inkludert i den endelige modellen fordi den ble regnet som en substansiell viktig faktor i prediksjonen av statsministerprestasjoner. I tillegg var den signifikant i bivariat regresjon og hadde p-verdi svært nær signifikansnivået i multivariat regresjon.

På grunnlag av de disse analysene endte besto den endelige modellen av variablene *antall måneder* og *arbeidsledighet*. Justert R^2 var på 52.9 prosent. I denne analysen var b-verdien til *antall måneder* 0.017. Ved å sitte fem år lenger ved makten oppnår en statsminister altså 1 score høyere vurdering. I et komparativt perspektiv er resultatet av variabelen *antall måneder* svært interessant. Simonton refererer til fire ulike studier der lengden på en presidents regjeringstid gir signifikant prediksjon på storhetsgrad (1986: 270). Dette indikerer også at lengden en statsminister og president blir sittende med makten er en signifikant predikator på storhet på tross av store institusjonelle ulikheter mellom landene. Selv om må her må ta i betraktning at storhetsgrad måles litt ulikt mellom analysene er funnet interessant. Dessuten indikerer *antall måneder* at velgerens oppfatning om hvilke kriterier som skal til for å bli vurdert som en god statsminister samsvarer med de respondentenes legger til grunn.

Den makroøkonomiske situasjonen i Norge har også betydning for hvordan statsministernes prestasjoner blir vurdert. En statsminister som regjerte under perioder med for eksempel åtte prosent arbeidsledighet får 0.5 lavere score enn en som regjerte med tre prosent ledighet. Man kan derfor argumentere for at stabilitet i landet og levestandarden til innbyggerne har betydning for storhetsgraden til norske statsministere. På tross av at arbeidsledigheten i større eller mindre grad avhenger av internasjonale konjunkturer, holder respondentene statsministeren delvis ansvarlig for den sosioøkonomiske utviklingen. Dette funnet støtter opp om kritikken av denne type rangeringer, det Ross kaller "the fundamental attribution error" (sitert i Simonton 1986: 279). Ved bare å fokusere på individ og dermed tillegge statsministere egenskaper de ikke selv har kontroll over, overser man de situasjonsbetingede begrensningene i deres handlemåte.

Residualanalysen viste at Christian Michelsen, John Lyng, Christopher Hornsrud, Wollert Konow, Thorbjørn Jagland og Odvar Nordli ble dårligst predikert av modellen. Førstnevnte

var også signifikant forskjellig fra de andre statsministerne. Michelsen kan også betegnes som et innflytelsesrikt case fordi han påvirker modellens samlede effekt uforholdsmessig mye. Einar Gerhardsen er med sine 209 måneder som statsminister et innflytelsesrikt case på variabelen *antall måneder*.

Ut fra den endelige modellen ble Kjell Magne Bondevik og Jens Stoltenbergs storhetsgrad predikert. De fikk scoren 3.81 og 3.71 og hadde dermed endt opp på tiende og ellefte plass over Norges største statsministere.

6.2 Hvor går veien videre?

Den endelige modellen viste at *antall måneder* og *arbeidsledighet* fungerer som sikre prediktorer på storhetsgraden til norske statsministere i perioden 1905-1997. Dette er funn som er verdifulle for fremtidig forskning på feltet. I den internasjonale litteraturen er sammenligninger mellom ulike studier og funn svært vesentlig. Grunnen til dette er at de gjør det mulig å forsterke antakelsene om hvilke variabler som predikerer/ikke predikerer storhet uavhengig av respondentene som vurderer.

Denne eksplorative studien kan være grunnlag for en utvikling av forskningsfeltet i Norge. En åpenbar videreføring av analysen er å foreta en ny "Cordtsen-undersøkelse" der respondenter vurderer statsministerprestasjoner. Med dette som grunnlag for verdiene på avhengig variabel, kan man deretter analysere hvilke faktorer som predikerer statsministerprestasjoner. I en ny studie må man være oppmerksom på en del punkter: Hvem bør delta? Det naturlige er å ha et utvalg bestående av historikere og statsvitere da personer innenfor denne gruppen ofte har best faglige forutsetninger for å delta. Det anbefales dessuten å inkludere andre respondenter enn dem som deltok i Cordtsen-studien slik at man kan identifisere eventuelle avvik mellom respondentenes vurderinger. Et sentralt punkt her er å få med kvinnelige respondenter i undersøkelsen. Hvorfor svarer respondentene som de gjør? Det kan vise seg at det er systematiske tendenser i respondentenes vurderinger. For å få klarlagt slike forhold er det en fordel om man får presisert respondentenes alder, politiske ståsted, profesjonsstatus og geografisk bosted.

Når det gjelder utformingen av undersøkelsen kan man la respondentene selv vurdere kriteriene for storhet, slik det også er gjort i denne studien. På denne måten oppnår man et bra utgangspunkt for sammenligning. På den andre siden kunne det vært interessant å be

respondentene vurdere hver statsministers prestasjon på ulike områder slik Ballard og Suedfeld gjorde i sin studie av canadiske statsministerprestasjoner (1988). Metoden ble også benyttet av Lønnstrøm og Kelly. De brukte 21 definerte kategorier som amerikanske presidenter skulle vurderes på fra en skala fra en til fem. De mente dette minket sjansen for at subjektive meninger påvirket storhetsscoren (2003: 626). Eksempler på kategorier er politisk lederskap samt nasjonal- og utenrikspolitikk.

Mulighetene er altså flere i videreføringen av forskningsfeltet. Denne analysen har lagt et grunnlag for fremtidige studier, men det er trolig en rekke faktorer som predikerer storhet. Hva disse er og hvordan de predikerer storhet vil være opp til kommende studier å få svar på.

7. LITTERATURLISTE

- Aardal, Bernt, Høstmark, Maria, Lagerstrøm, Bengt Oscar, and Stavn, Guro. 2007. Valgundersøkelsen 2005, Statistisk Sentralbyrå
- Adler, David G. 2003. "Presidential Greatness as an Attribute of Warmaking" in *Presidential Studies Quarterly* 33(3):466-484
- Ballard, Elizabeth J., and Suedfeld, Peter. 1988. "Performance Ratings of Canadian Prime Ministers: Individual and Situational Factors" in *Political Psychology* 9(2 (June)):291-302
- Berntsen, Thorbjørn. 1997. Rottekjøret mot Thorbjørn Jagland. *Verdens Gang*
- Blessing, Tim H. 1995. "Three swings and ten dynasties: The temporal-historical aspects of presidential performance" in *Presidential Studies Quarterly* 25(3):457-469
- Bohrnstedt, George W., and Knoke, David. 1994. *Statistics for Social Data Analysis*. Itasca, Illinois: F.E.Peacock
- Bond, Jon R., Fleisher, Richard , and Krutz, Glen S. . 1996. *An Overview of the Empirical Findings on Presidential-Congressional Relations. In Rivals for Power. Presidential-Congressional Relations*. Washington D.C. : CQ Press
- Bose, Meena , and Landis, Mark. 2003. *The Uses and Abuses of Presidential Ratings*. New York: Nova Science Publishers, INC. 127 pp.
- Brende, Børge , Clemet, Kristin, Norvik, Erling, and Stray, Sverre. 1993. *Politikk og ansvar: Debattskrift til Kåre Willoch 3.oktober 1993*. Oslo: Schibsted. 232 pp.
- Brøgger, Waldemar m.fl. . 1983. *Cappelens leksikon*. Oslo: Cappelens forlag. 379 pp.
- Bull, Edvard. 1979. *Norges historie , bind 13: Klassekamp og fellesskap 1920-1945*: J.W. Cappelens Forlag
- Chatfield, Chris. 1996. *The Analysis of Time Series: An Introduction*. London: Chapman & Hall. 283 pp.
- Cohen, Jeffrey E. 2004. "The polls: Presidential Greatness as Seen in the Mass Public: An Extension and Application of the Simonton Model " in *Presidential Studies Quarterly* 33(4):913-945
- Collier, David , and Mahon, James E.jr. 1993. "Conceptual "Stretching" Revisited: Adapting Categories in Comparative Analysis" in *The American Political Science Review* 87(4 (December 1993)):845-855
- Cordtsen, Kjell. 1997a. En historisk vurdering av Norges statsministere. *Verdens Gang*
- Cordtsen, Kjell. 1997b. Gerhardsen til topps, Jagland til bunns. *Verdens Gang*:21-25
- Crockett, D.A. 2002. *The Opposition Presidency: Leadership and the Constraints of History*: Texas A&M University Press. 320 pp.
- Furre, Berge. 1991. *Vårt Hundreår Norsk Historie 1905-1990*. Oslo: Det Norske Samlaget. 561 pp.
- Furre, Berge. 2000. *Norsk Historie 1914-2000*. Oslo: Det Norske Samlaget
- Gelman, Andrew, and Stern, Hal. 2006. "The Difference Between "Significant" and "Not Significant" is not Itself Statistically Significant" in *The American Statistician* 60(4):328-331
- Grimen, Harald. 2000. *Samfunnsvitenskapelige tenkemåter*. Oslo: Universitetsforlaget. 424 pp.
- Grytten, Frode. 1994. *En empirisk analyse av det norske arbeidsmarkedet 1918-1939 Arbeidsledigheten i Norge i et internasjonalt perspektiv* ph.d Bergen. 392 pp.
- Gujarati, Damodar. 1995. *Basic Econometrics*. London: McGraw Hill
- Hair, Joseph F., Anderson, Rolph E., Tatham, Ronald L., and Black, William C. 1998. *Multivariate Data Analysis*. New Jersey: Prentice- Hall PTR. 729 pp.

- Hair, Joseph F. jr., Black, William C, Babin, Barry J., Anderson, Rolph E., and Tatham, Ronald L. 2006. *Multivariate Data Analysis*. New Jersey: Pearson Prentice Hall. 881 pp.
- Hansson, Steinar, and Teigene, Ingolf Håkon. 1992. *Makt og Mannefall- Historien om Gro Harlem Brundtland*. Oslo: Cappelen
- Heidar, Knut, and Berntzen, Einar. 1998. *Vesteuropeisk politikk Partier, regjeringsmakt, styreform*. Oslo: Universitetsforlaget. 375 pp.
- Hellevik, Ottar. 2002. *Forskningsmetode i sosiologi og statsvitenskap*. Oslo: Universitetsforlaget. 487 pp.
- Johannsesen, Asbjørn, and Tuft, Per Arne. 2002. *Introduksjon til samfunnsvitenskapelig metode*. Oslo: Abstrakt forlag. 275 pp.
- Keilhau, Wilhelm. 1927. *Norge og Verdenskrigene*. Oslo: Aschehough og co
- Kjølås, Harald. 1999. *Norske Statsministre* Oslo: Det Norske Samlaget. 355 pp.
- Krueger, James S., and Lewis-Beck, Michael S. 2007. "Goodness-of-Fit: R-Squared, SEE and "best practice"" in *The Political Methodologist* 15(1):1-4
- Lahlum, Hans Olav. 2007. *Oscar Torp- en politisk biografi*. Oslo: Cappelen 560 pp.
- Lewis-Beck, Michael S. 1980. *Applied Regression: An Introduction*: Sage Publications. 74 pp.
- Lijphart, Arend 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven and London: Yale University Press. 368 pp.
- Linz, Juan J. 1990. "The Perils of Presidentialism" in *Journal of Democracy* 1 (winter):51-69
- Lund, Jørn. 2000. *Den Store Danske Encyklopædi*. København: Nordisk Forlag A/S. 654 pp.
- Lønnstrøm, Douglas A., and Kelly, Thomas O. 2003. "The Contemporary Presidency: Rating the Presidents: A Tracking Study" in *Presidential Studies Quarterly* 33(3 (September)):625-634
- Maranell, Gary M. 1970. "The Evaluation of : An Extension of the Schlesinger Polls" in *The Journal of American History* 57(1 (June)):104-113
- Midtbø, Tor. 2007. *Regresjonsanalyse for samfunnsvitere Med eksempler fra SPSS*. Oslo: Universitetsforlaget. 131 pp.
- Murray, Robert K., and Blessing, Tim H. 1994. *Greatness in the White House: Rating the Presidents* Penn State Press
- Murray, Robert K., and Blessing, Tim H. . 1983. "The Presidential Performance Study: A Progress Report" in *The Journal of American History* 70(No.3 (December)):535-555
- Nice, David C. 1984. "The Influence of War and Party System Aging on the Ranking of Presidents" in *The Western Political Quarterly* 37(3 (September)):443-455
- Olstad, Finn. 1999. *Einar Gerhardsen- en politisk biografi*. Oslo: Universitetsforlaget. 479 pp.
- Pedhazur, Elazar J. 1997. *Multippel Regression in Behavioral Research: Explanation and Prediction*: Thomson Learning. 1058 pp.
- Pincus, Walter. 2008. Eisenhower Advisers Discussed Using Nuclear Weapons in China. *The Washington Post*
- Ragin, Charles C. 1987. *The Comparative Method Moving beyond qualitative and Quantitative Strategies*. California: University of California Press. 185 pp.
- Ragin, Charles C., and Zaret, David. 1983. "Theory and Method in Comparative Research: Two Strategies" in *Social Forces* 61(3):731-754
- Ranson, K.Anne. 2005. *Encyclopedia AMERICANA* Danbury, Connecticut: Scholastic Library Publishing, Inc. 879 pp.

- Rokkan, S. 1966. "Norway: Numerical Democracy and Corporate Pluralism". In *Political Opposition in Western Democracies*, ed. Dahl, Robert. A., pp. 70-115: Yale University Press
- Rokkan, S. 1987. *Stat, nasjon, klasse*: Universitetsforlaget 403 pp.
- Schlesinger, Arthur M. 1979. "Rating the Presidents: Washington to Clinton" in Schlesinger jr., Arthur M. 1997. "Rating the Presidents: Washington to Clinton" in *Political Science Quarterly* 112(2 (summer)):179-190
- Schlesinger jr., Arthur M. . 1996. The Ultimate Approval Rating. *The New York Times Magazine*:45-51
- Simonton, Dean Keith. 1986. "Presidential Greatness: The Historical Consensus and Its Psychological Significance" in *Political Psychology* 7(2 (June)):259-283
- Simonton, Dean Keith. 2001. "Predicting Presidential performance in the United States: Equation replication on recent survey results" in *The Journal of Social Psychology* 141(3):293-301
- Skog, Ole Jørgen. 1998. *Å forklare sosiale fenomener: en regresjonsbasert tilnærming* Oslo: Ad Notam Gyldendal. 329 pp.
- Skog, Ole Jørgen. 2004. *Å forklare sosiale fenomener: en regresjonsbasert tilnærming*. Oslo: Gyldendal akademisk
- Slagstad, Rune. 1998. *De nasjonale strateger*. Oslo: Pax Forlag A/S. 561 pp.
- Steenbergen, Marco R., and Marks, Gary. 2007. "Evaluating expert judgments" in *European Journal of Political Research* 46(3):347-366
- Taranto, James. 2004. "What Makes a President Great? Scholars finally begin giving Reagan his due" in *Opinion Journal* (Thursday, June 10):1-3
- Taranto, James. 2005. "How is He Doing? George W. Bush is "average," but far from ordinary" in *Opinion Journal* (Monday, September 12):2
- Theakston, Kevin. 2003. Political skills and context in prime ministerial leadership in Britain. In *Leadership in Context*, ed. Hargrove, E.C., and Owens, J.E., pp. 89-129. London: Rowman & Littlefield
- Theakston, Kevin , and Gill, Mark. 2006. "Rating 20th- Century British Prime Ministers " in *The British Journal of Politics and International Relations* 8(2):198-213
- Winter, D.G. 1987. "Leader appeal, leader performance, and the motive profiles of leaders and followers: A study of American presidents and elections." in *Journal of Personality and Social Psychology* 52:196-202
- Wyller, Thomas Chr. 1975. *Christian Michelsen- politikeren*. Oslo: Oslo Forlagstrykkeri AS. 250 pp.
- Østerud, Øyvind, Engelstad, Fredrik, and Selle, Per. 2003. *Makten og demokratiet En sluttbok fra Makt- og demokratiutredningen*. Oslo: Gyldendal Norsk Forlag AS. 344 pp.

8. APPENDIKS

Appendiks A – variabeloversikt

(for utfyllende variabelnavn se figur 1 side 16)

Statsminister	Nyttpart	Person	Koalisjo	Flertall	Antmåned	Perioder	Polsyst	Gjenvalg	Konfront	Strid
Christian Michelsen	1	0	1	1	32	0	0	0	0	1
Jørgen Løvland	0	0	1	1	5	0	0	0	1	0
Gunnar Knudsen	1	0	0	1	114	1	0	1	1	0
Wollert Konows	1	0	1	1	25	0	0	0	0	1
Jens Bratlie	1	0	1	1	11	0	0	1	0	0
Otto B. Halvorsen	1	0	1	0	16	1	0	0	1	0
Otto A. Blehr	1	0	0	0	21	0	0	1	1	0
Abraham Berge	1	1	1	0	14	0	0	0	1	0
Johan L. Mowinckel	1	0	0	0	86	1	0	1	1	0
Ivar Lykke	1	0	1	0	23	0	0	1	0	0
Christopher Hornsrud	1	0	0	0	1	0	1	0	1	0
Peder L. Kolstad	1	0	0	0	11	0	1	0	0	0
Jens Hundseid	0	1	0	0	12	0	1	0	1	0
Johan Nygaardsvold	1	0	0	0	124	0	1	1	0	0
Einar Gerhardsen	0	0	0	1	209	1	1	1	1	0
Oscar Torp	0	1	0	1	39	0	1	1	0	1
John Lyng	1	0	1	0	1	0	1	0	1	0
Per Borten	1	0	1	1	66	0	1	1	0	1
Trygve Bratteli	1	0	0	0	48	1	1	0	0	1
Lars Korvald	1	0	1	0	12	0	1	0	0	0
Odvar Nordli	0	0	0	0	61	0	1	1	0	0
Gro H. Brundtland	0	1	0	0	123	1	1	1	0	0
Kåre Willoch	1	0	1	1	56	1	1	1	1	0
Jan P. Syse	1	0	1	0	13	0	1	0	0	1
Thorbjørn Jagland	0	1	0	0	12	0	1	0	0	0

Statsminister	Valg	Bnpendri	Arbledig	Utdannin	Region	Alderinn	Leveår	Sdffit	Sdfbeta ant måned	Sdfbeta arbledig
Christian Michelsen	0	2.9	1.1	2	1	47	17	0.97	-0.29	-0.61
Jørgen Løvland	0	2.6	1.2	1	2	59	14	-0.08	0.05	0.04
Gunnar Knudsen	0	1.5	1.1	2	2	59	8	0.25	0.18	-0.08
Wollert Konows	0	1.9	0.9	1	1	64	12	-0.49	0.19	0.32
Jens Bratlie	1	3.5	0.6	2	2	56	19	-0.33	0.17	0.22
Otto B. Halvorsen	0	-0.2	1.7	2	2	48	0	-0.06	0.03	0.03
Otto A. Blehr	0	5.4	6.6	2	1	74	14	0.00	0.00	0.00
Abraham Berge	0	0.0	4.5	1	1	71	12	-0.14	0.07	-0.02
Johan L. Mowinckel	0	3.8	8.3	2	1	53	8	0.19	0.10	0.15
Ivar Lykke	1	1.9	8.1	1	0	54	20	-0.11	0.01	-0.08
Christopher Hornsrud	0	4.8	7.3	0	2	68	29	0.64	-0.27	0.39
Peder L. Kolstad	0	-1.1	10.4	1	2	52	0	-0.15	0.02	-0.13
Jens Hundseid	0	2.5	9.8	1	2	48	30	-0.40	0.07	-0.34
Johan Nygaardsvold	0	2.5	7.1	0	0	55	6	0.14	0.11	0.08
Einar Gerhardsen	0	4.3	0.9	1	2	48	23	-0.75	-0.69	0.08
Oscar Torp	0	3.1	0.9	0	2	58	3	-0.14	0.03	0.09
John Lyng	0	3.0	1.3	2	0	58	14	0.60	-0.38	-0.32
Per Borten	0	3.2	0.9	0	0	52	33	-0.13	-0.03	0.08
Trygve Bratteli	0	2.6	0.9	1	2	61	8	0.08	0.00	-0.05
Lars Korvald	0	2.9	0.8	1	2	56	33	0.06	-0.03	-0.04
Odvar Nordli	0	3.6	1.1	1	2	48	27	-0.28	-0.04	0.16
Gro H. Brundtland	1	2.2	3.0	2	2	41	25	0.13	0.11	0.00
Kåre Willoch	0	3.5	2.6	2	2	53	21	0.21	0.03	-0.05
Jan P. Syse	0	1.7	4.3	2	2	58	19	-0.13	0.07	-0.01
Thorbjørn Jagland	1	2.7	4.0	2	2	45	32	-0.29	0.16	-0.01

- Utdanning er kodet på følgende måte: 0= grunnskole, 1= over grunnskole og 2= Hovedfag.
- Region er kodet på følgende måte: 0=Østlandet, 1= Sør-Vest Norge og 2= Nord-Norge.

Appendiks B – respondentene i Cordtsen-studien

Berntsen,Harald	historiker og biografiforfatter
Dahl,Hans F.	historiker og professor i media og kommunikasjon
Dahl,Svein	Historiker
Danielsen,Rolf	tidligere professor i historie ved UiB
Eriksen,Knut E.	professor 11 ved UiO
Hagtvet,Bernt	professor i statsvitenskap
Heidar,Knut	professor i statsvitenskap og dekan ved Institutt for samfunnsvitenskap, UiO
Kokkvold,Arne	historiker
Lundestad,Geir	professor i historie og leder for Det Norske Nobelinstitutt
Moland,Arnfinn	førsteamanuensis i historie og direktør ved Norges Hjemmefrontmuseum
Nordby,Trond	historiker og professor i statsvitenskap ved UiO
Norvik,Per	politisk kommentator og historiker
Olstad,Finn	forfatter og historiker
Pharo,Helge	professor i historie ved UiO
Pryser,Tore	professor i historie ved HiL
Riste,Olav	professor emeritus og forskningssjef ved Institutt for forsvarsstudier
Ringdal,Nils J.	historiker
Solstad,Arve	statsviter og tidligere redaktør i Dagbladet
Sørensen,Øystein	professor i historie ved UiO
Versto,Olav	politisk redaktør og historiker

* Stillingsbeskrivelsene bygger på offentlig tilgjengelig informasjon, men det tas høyde for at respondentene kan ha andre/nye roller enn dem beskrevet her.

Appendiks C – korrelasjonsmatrise

	Nyttpart	Person	Koalisjo	Flertall	Antmån	Antstats	Polsys	Gjenvalg
Nyttpart	1.000							
Person	-0.579**	1.000						
Koalisjo	0.421*	-0.280	1.000					
Flertall	-0.089	-0.167	0.280	1.000				
Antmån	-0.256	-0.054	-0.435*	0.248	1.000			
Antstats	-0.008	-0.089	-0.243	0.089	0.598**	1.000		
Polsys	-0.327	0,204	-0.360	-0.238	0.175	-0.036	1.000	
Gjenvalg	-0.114	-0.080	-0.282	0.280	0.624**	0.370	0.292	1.000

	Nyttpart	person	Koalisjo	Flertall	Antmån	Antstats	Polsys	Gjenvalg
Konfront	0.014	-0.400	-0.450	0.007	0.058	0.345	-0.263	-0.045
Strid	0.142	-0.470	0.210	0.359	-0.093	-0.142	0.076	-0.165
Valg	-0.214	0.327	0.017	-0.100	-0.028	-0.029	-0.089	0.236
BNP	-0.177	-0.170	-0.232	0.184	0.205	-0.030	0.147	0.412*
Arbledig	0.116	0.136	-0.376	-0.578**	-0.150	-0.183	0.043	-0.043
Utdannin	0.118	-0.055	0.182	-0.060	-0.070	0.375	-0.362	-0.040
Region	-0.430*	0.212	-0.293	0.075	-0.170	0.312	0.195	-0.187
Alder	0.466*	-0.181	0.109	-0.031	-0.349	-0.284	-0.318	-0.146
Leveår	-0.307	0.166	0.072	-0.031	-0.015	-0.237	0.382	0.016

	Konfront	Strid	Valg	BNP	Arbledig	Utdannin
Konfront	1.000					
Strid	-0.498*	1.000				
Valg	-0.387	-0.245	1.000			
BNPendr	0.151	-0.012	-0.005	1.000		
Arbledig	0.152	-0.368	0.048	-0.139	1.000	
Utdannin	0.214	-0.218	0.284	-0.660	-0.116	1.000
Region	0.077	-0.109	0.012	-0.056	-0.162	0.194
Alder	0.299	0.088	-0.358	0.141	0.076	-0.205
Leveår	-0.120	-0.098	0.302	0.473*	-0.080	-0.075

	Region	Alder	Leveår
Region	1.000		
Alder	-0.177	1.000	
Leveår	0.069	-0.251	1.000

* Variablene i kolonne 1 har samme rekkefølge som variabelpresentasjonen i figur 1 på side 16.

Appendiks D – kronologisk rangering av Norges statsministere i perioden 1905-1997

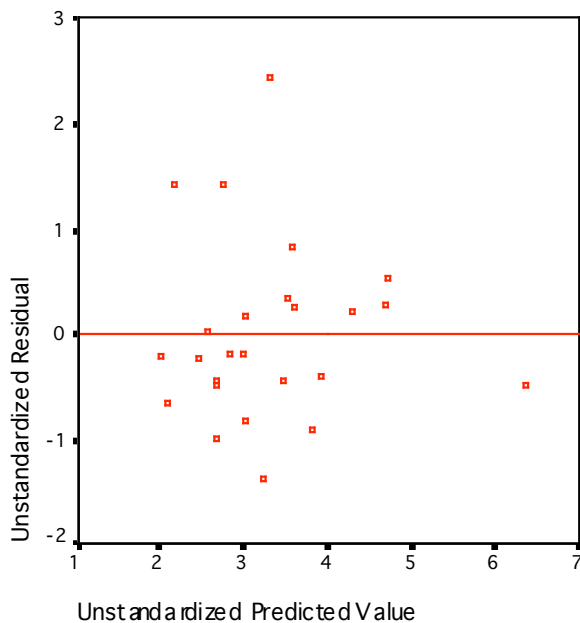
Statsministere	storhetsscore	Rangering
Einar Gerhardsen	5.85	1
Christian Michelsen	5.75	2
Gunnar Knudsen	5.25	3
Gro H.Brundtland	4.95	4
Johan Nygaardsvold	4.50	5
Kåre Willoch	4.40	6
John Lyng	4.18	7
Johan L.Mowinckel	3.85	8
Trygve Bratteli	3.85	8
Christopher Hornsrud	3.56	10
Per Borten	3.50	11
Lars Korvald	3.16	12
Oscar Torp	3.00	13
Odvar Nordli	2.90	14
Otto B.Halvorsen	2.78	15
Jørgen Løvland	2.63	16
Otto Blehr	2.58	17
Ivar Lykke	2.21	18
Jan P.Syse	2.21	18
Jens Bratlie	2.18	20
Abraham Berge	2.17	21
Wollert Konow	1.84	22
Peder Kolstad	1.80	23
Thorbjørn Jagland	1.68	24
Jens Hunseid	1.45	25

Appendiks E – test av forutsetninger

Vanlige antakelser for OLS- regresjon er 1) ingen heteroskedastisitet og 2) normalfordelte restledd. I tillegg til dette forutsettes det 3) ingen autokorrelasjon eller 4) multikollinearitet. De to sistnevnte ble testet for ved hjelp av henholdsvis Durbin-Watson og toleranseverdier i analysens ulike modeller.

Homoskedastisitet refererer til sammenhengen mellom spredningen av residualene og de predikerte uavhengige variablene. Det forutsettes at det skal være lik feilvarians over hele avhengig variabel for ulike verdier av de uavhengige variablene. Hvis residualene øker eller minker med verdiene av de uavhengige variablene oppstår heteroskedastisitet, og forutsetningen er brutt . Test av homoskedastisitet er vist under.

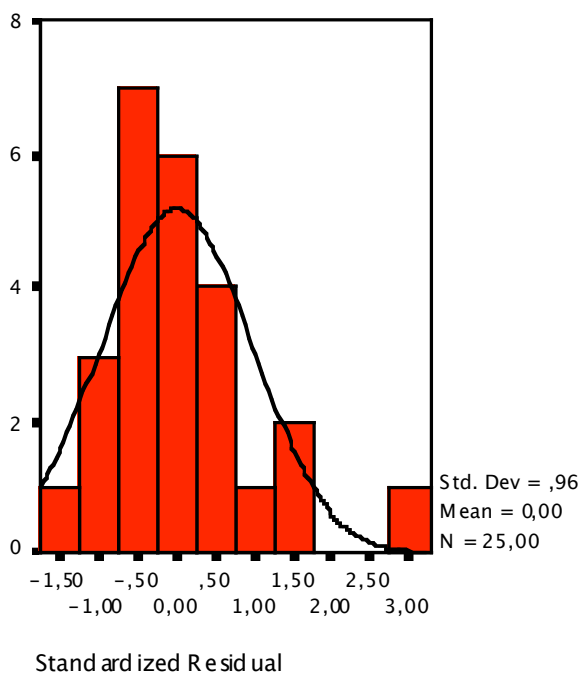
Residualer og predikerte verdier



På tross av at det ikke er perfekt homoskedastisitet er det ikke tendens til mønster i verdiene, og det er således ikke et brudd på forutsetningen om homoskedastisitet.

Normalfordelte restleddverdier. Ekstremverdier kan føre til at residualene ikke er tilstrekkelig normalfordelt. Dette kan videre gjøre signifikanstesting upålitelig. De standardiserte residualverdiene er vist under.

Test av normalfordelte residualer



Figuren viser at residualfordelingen er høyreskjev, og årsaken til dette er Christian Michelsen som har ekstremverdi. Verdiene for dffit og dfbeta bør ligge mellom -2 og 2, og Michelsen har standardisert residualverdi på 2.75. Årsakene til dette er allerede diskutert i dette kapittelet. Ser man bort fra dette viser figuren at restleddene i stor grad er normalfordelte.