

Sammanfattning

Varför regerar vissa ledare längre än andra? Uppsatsen använder ett datamaterial över ledares tid vid makten i 167 länder från början av 1800-talet till 1987 för att försöka besvara denna fråga. Datamaterialet bearbetas med olika överlevnadsanalysmetoder för att kunna fastställa vilken modell som ger det mest tillförlitliga och informativa resultatet. Bland annat används en Cox regressionsmodell, en accelerated failure time modell och en multinomial logit modell. Modellernas antaganden testas och resultaten utvärderas därefter. Ledarens tid vid makten modelleras med konstitutionell och icke konstitutionell maktförlust som händelse. Kumulativ incidens används för att ta hänsyn till att händelserna kan vara beroende gentemot varandra. Resultatet tyder på att en Cox regressionsmodell verkar vara den mest robusta modellen. Det är även mer intressant att modellera händelserna var för sig (än tillsammans) eftersom variablerna påverkar händelserna mycket olika. Beroende mellan konkurrerande risker kan dock orsaka viss bias. Vad gäller variablerna är regionen ledaren styr i den som oftast har störst effekt på risk för maktförlust och överlevnadstid. Detta kan dock bero på att variabeln agerar som en sammanfattning av flera svåråtkämpliga sociala, ekonomiska och kulturella aspekter som grannländer ofta delar med varandra.

Förord

Vi vill tacka vår handledare Gebrenegus Ghilagaber för att ha låtit oss ta del av det datamaterial som utgör grunden i uppsatsen och den engagerande handledningen. Uppsatsen har skrivits gemensamt av författarna och kritik förväntas ges till författarna ömsesidigt.

Innehållsförteckning

1 Inledning	4
1.1 Syfte och frågeställning	4
1.2 Tidigare forskning	5
2 Metod	6
2.1 Överlevnadsanalys	6
2.1.1 Kaplan-Meier estimatorn av överlevnadsfunktionen	7
2.1.2 Life-table estimatorn av hazardfunktionen	7
2.1.3 Log-rank och Wilcoxon	8
2.2 Parametrisk modell	9
2.2.1 Fördelningar	9
2.2.2 Goodness of fit test (likelihood-ratio)	12
2.3 Semiparametrisk modell	13
2.3.1 Test av proportionaliteten i Cox-modellen	14
2.4 Multinomial logit-modell	15
2.5 Kumulativ incidens	16
2.5.1 Pepe och Moris metod	17
3 Beskrivning av data	18
3.1 Variablerna	18
3.2 Ledarnas överlevnad för hela datamaterialet	18
3.3 Ledarnas överlevnad i regionerna	20
3.4 Konkurrerande risker	21
3.5 Vänstercensurering	22
3.6 Naturlig död	22
4 Resultat	23
4.1 Icke parametriska test för variablerna	23
4.2 Semiparametrisk modell	23
4.2.1 Test av proportionalitetsantagandet	23
4.2.2 Anpassning av Cox regressionsmodell	25
4.3 Parametrisk modell	26
4.3.1 Distributionstest	26
4.3.2 Anpassning av AFT modell	27
4.4 Multinomial logit modell	29
4.5 Kumulativ incidens	30
5 Analys och diskussion	31
5.1 Variablerna	31
5.2 Bristfälliga metoder för konkurrerande risker	34
5.3 Vilken modell?	35
6 Slutord	36
Referenser	38
Appendix A - Variabelbeskrivning	39
Appendix B – Sammandrag av datamaterialet	40
Appendix C Sammandrag av resultat	44
Appendix D Modeller över subgrupper	50

1 Inledning

Nationer har länge styrts av regenter med olika karaktärer. De kan ha absolut makt eller vara folkvalda. Monarker och diktatorer sitter antingen säkert under hela sina liv på sina platser eller störtas och ersätts med något nytt. Ledare kan bli framröstade av folket och sedan utbyttas om de inte tycks vara dugliga eller när tiden är mogen för en ny ledare att ta över. En ledare kan förlora makten på många olika sätt och med minst lika många orsaker. Frågan är om det finns något mönster som kan förklara ledarnas tid vid makten och hur de sedan förlorar den. Varför sitter vissa ledare längre vid makten än andra? Vissa ledare sitter nästan en hel livslängd vid makten medan andra knappt hinner börja regera innan de avsätts. Är det ledarens personliga egenskaper som avgör eller är det landets förutsättningar med avseende på demokrati, kultur och tradition? Vad är det egentligen som bestämmer om en ledare ska få sitta kvar eller avsättas på ett eller annat sätt?

Detta är ständigt aktuella frågor som många statsvetare vill besvara. Henry Bienen och Nicolas van de Walle gjorde ett försök i sin kvantitativa studie år 1991. De modellerade ledares tid vid makten baserat på ett ambitiöst insamlat datamaterial som inkluderade ledare från tidigt 1800-tal till 1987 i 167 länder samt mer än 20 olika variabler. De reella ledarna i varje land identifierades för olika tidpunkter och hänsyn togs till att en ledare kunde avsättas på flera olika sätt. I studien finns tre typer av maktförluster; konstitutionell, icke konstitutionell och att ledare dör av naturliga orsaker. Dessutom samlades variabler in som antogs kunna förklara varför tiden vid makten (överlevnadstiden) hade en specifik längd.

Detta datamaterial är grunden för denna uppsats som med fler tillgängliga statistiska modeller kommer att ge en vidare förståelse i ämnet, samtidigt som metoderna utvärderas utifrån förmågan att modellera datamaterialet.

1.1 Syfte och frågeställning

Bienen och van de Walle (1991) har i sin studie använt överlevnadsanalys för att modellera ledares tid vid makten och efterföljande maktförlust. Det insamlade datamaterialet är stort och ambitiöst men den statistiska analysen är mest deskriptiv och de enda modellerna som anpassas är en semiparametrisk proportionell hazard modell och en tidsberoende version (som är flera modeller för olika tidsintervall av överlevnadstiden). Inom överlevnadsanalysen finns det flera andra metoder som möjligtvis skulle kunna förklara datamaterialet bättre än just dessa modeller. Denna uppsats har till syfte att undersöka om det kan finnas en eller flera metoder/modeller som mer korrekt kan hjälpa till att besvara Bienen och van de Walles frågeställningar. Där korrekt avser att även testa metodernas antaganden för att kontrollera att de är uppfyllda, någonting som Bienen och van de Walle bortser ifrån. Författarna modellerade ej heller händelserna var för sig, vilket denna uppsats avser att göra. Händelserna är ömsesidigt uteslutande, vilket betyder att de är *konkurrerande risker* gentemot varandra. Då en maktförlust skett, resulterar det i att inga andra händelser kan inträffa.

Huvudfrågeställningen i Bienen och van de Walles studie är varför vissa ledare är kvar vid makten längre än andra. Till detta har författarna även några hypoteser och problem som de försöker besvara genom att inkludera olika variabler i deras modeller. Nedanför följer några frågeställningar som denna uppsats ska försöka besvara. Valet av dessa frågor är baserade på att, givet frågornas svar, nå upp till uppsatsens syfte att finna bättre modeller som beskriver datamaterialets egenskaper på ett korrekt vis.

- ❖ Hur är datamaterialet fördelat? Följer ledarnas överlevnad någon specifik distribution?
- ❖ Kan en modell med *proportionell* hazard antas för datamaterialet? Om inte, hur kan detta då hanteras?
- ❖ Hur förklarar variablerna ledarnas överlevnad? Är resultatet som väntat? Om inte, hur ska detta tolkas?
- ❖ Hur skiljer sig händelserna (icke konstitutionell och konstitutionell maktförlust) sinsemellan? Har variablerna olika effekter för vardera händelse? Påverkar detta i så fall valet av metod?
- ❖ Finns det beroende mellan händelserna? Vilka konsekvenser för det i så fall med sig för de valda modellerna och hur kan det hanteras?

I denna uppsats kommer konstitutionell och icke konstitutionell maktförlust att analyseras. Den tredje händelsen, naturlig död, kommer inte analyseras. Detta beror på uppsatsens syfte att analysera ledare som förlorar makten genom att bli avsatt, vilket inte är fallet för naturlig död som även kan antas påverkas av helt andra variabler än de två tidigare. Det är mycket få observationer med ledare som förlorar makten genom att naturligt avlida vilket ger ett ytterliggare skäl till att inte inkludera dessa i analysen. En vidare argumentation för detta kommer senare i uppsatsen. Även vissa variabler kommer ges mindre utrymme i uppsatsen då dessa endast är aktuella för vissa regioner och specifika tidsperioder.

1.2 Tidigare forskning

Bienen och van de Walles slutsats av studien är att den bästa modellen är en proportionell hazard modell. Vidare drogs slutsatsen att hazarden inte är konstant utan snarare minskar risken att förlora makten över tiden. Ju längre en ledare sitter vid makten, ju mindre risk är det att han eller hon kommer att förlora den. Vad detta sedan beror på var inte deras studie kapabel att besvara. Andra slutsatser vad gällde variablerna var att nackdelen att ta makten icke-konstitutionellt försvinner efter några år. De är tydliga med att poängtera att variabeln för typ av makt tillträde är tidsberoende, detta betyder i sin tur att ett av grundantagandena för en proportionell hazard modell inte är uppfyllt, men modellen anses trots det som den bästa.

Utöver Bienen och van de Walle själva har även Paul Allison (1995) gjort en studie av en mindre del av datamaterialet. Detta omfattade ledare som regerade från och med år 1960 fram till studiens slutår 1987. Allison exkluderade ledare från Nordamerika, Europa och Australien samt observationer av ledare som regerar sin andra eller tredje gång och så vidare. Han introducerar fenomenet om konkurrerande risker och använder fler metoder på datamaterialet än vad Bienen och van de Walle (1991) gör. Varje händelse modelleras var för sig samt tillsammans som gemensam händelse. De metoder som används är Cox regressionsmodell, accelerated failure time-modell samt slutligen en multinomial logit-modell där han jämför konstitutionell och icke-konstitutionell maktförlust. Olika distributioner anpassas och testas och problemet med händelser som är informativa gentemot varandra diskuteras men inget åtgärdas. Slutsatsen är att den bästa modellen för materialet är en Cox regressionsmodell stratifierat efter region för konstitutionell och icke-konstitutionell maktförlust som gemensam händelse medan naturlig död censureras. Även logit-modellen ansågs representera aspekterna i datamaterialet på ett adekvat sätt. Slutligen konstaterar Allison att trots problem med

händelser som är informativa mot varandra så finns det inga utvecklade modeller baserade på beroendeförhållanden som kan ersätta vedertagna modeller som Cox och accelerated failure time för analyser av konkurrerande risker.

År 2006 kom dock Melania Pintilie ut med en bok där hon ger ett praktiskt perspektiv på hur konkurrerande risker kan hanteras. En modell för kumulativa incidenser ges där flera händelser kan inträffa. Det finns även en betingad sannolikhetsmodell för en händelse av intresse, där hänsyn tas till när den konkurrerande risken förändras. Utifrån resultat av Pepe och Mori (1993) ges också en metod för att testa om olika variabelvärden signifikant skiljer sig från varandra för kumulativa incidenser och betingade sannolikheter. Denna metod tar hänsyn till att händelserna är informativa och kan påverka varandra medan exempelvis Cox modellerar ”rena” effekter av kovariaterna för en händelse där andra händelser antas vara icke informativa.

I del 2 kommer metoderna som används i uppsatsen att beskrivas. Sedan kommer datamaterialets egenskaper presenteras och förklaras i del 3. Anledningen till att metoderna presenteras först är för att dessa är nödvändiga för presentationen av datamaterialet. Efter det redovisas resultatet i del 4 där de olika metoderna tillämpas på datamaterialets variabler. I uppsatsen kommer Allison's tillvägagångssätt med hänsyn till konkurrerande risker att tillämpas och även breddas. Hela datamaterialet analyseras samt konstitutionell och icke-konstitutionell maktförlust var för sig. Olika distributioner testas och den med bäst anpassning kommer att användas för modellering med variablerna. Cox regressionsmodeller kommer att användas med ytterliggare metoder för att testa proportionalitetsantagandet. En vidare problematisering kommer att ske angående beroendet mellan händelserna och metoder baserade på kumulativa incidenser används för att testa variablerna. Sedan diskuteras och analyseras resultatet i del 5 och så presenteras uppsatsens slutsatser. Skattningar kommer att jämföras med liknande metoder som ej tar hänsyn till konkurrerande risker. Slutligen kommer slutsatser att dras angående för- och nackdelar med metoderna, vilket som är det mest lämpliga modellvalet och vad som behöver uträttas för att kunna få ännu bättre skattningar för datamaterialet.

2 Metod

2.1 Överlevnadsanalys

När det analyseras hur lång tid det tar tills en specifik händelse inträffar så är överlevnadsanalys ett utmärkt redskap. För att summera sådan typ av data är det två olika funktioner som är av intresse och det är *överlevnadsfunktionen* och *hazardfunktionen*. Överlevnadstiden för en individ, t , kan ses som ett värde på en variabel, T , och kan ta vilket icke negativt värde som helst. Om den slumpmässiga variabeln T antas ha en sannolikhetsfördelning med den underliggande *täthetsfunktionen* $f(t)$, då fås en *Överlevnadsfunktion*, $S(t)$, som ser ut som följer

$$S(t) = P(T \geq t) = 1 - P(T < t) = 1 - \left[\int_0^t f(u) du \right]$$

som beskriver sannolikheten att överlevnadstiden är större eller lika med t (Collett, 2003). *Hazardfunktionen* används däremot för att beskriva risken för att uppleva en specifik händelse

för någon tidpunkt t , givet att individen har överlevt tills den tidpunkten. Om vi antar att en individs överlevnadstid T ligger mellan t och $t + \delta t$ så ser hazardfunktionen ut på följande vis

$$h(t) = \lim_{\delta t \rightarrow 0} \left[\frac{P(t \leq T < t + \delta t \mid T \geq t)}{\delta t} \right].$$

Ett annat samband är att

$$S(t) = \exp[-H(t)],$$

där $H(t)$ är integralen av hazarden $h(t)$. Funktionen $H(t)$ brukar kallas för den kumulativa hazarden (Collett, 2003). $H(t)$ kan även tas via överlevnadsfunktionen genom att den logaritmeras. Detta samband mellan den kumulativa hazarden och överlevnadsfunktionen kommer bland annat användas för att testa proportionalitetsantagandet i Cox proportionella hazard. Dock behövs metoder för att skatta överlevnaden och hazarden.

2.1.1 Kaplan-Meier estimatorn av överlevnadsfunktionen

För att analysera censurerad och icke grupperad överlevnadsdata används i normala fall estimation med Kaplan-Meier (Collett, 2003). För varje tidpunkt t finns det n_j individer som är utsatta för risk för en händelse. Med detta menas att individerna inte har upplevt en händelse eller blivit censurerade före tidpunkten t . Om δ_j är antalet individer som upplever händelsen av intresse under tidpunkt t då kommer Kaplan-Meier estimatorn se ut som följer

$$\hat{S}(t) = \prod_{j=1}^k \frac{n_j - \delta_j}{n_j}$$

för $t_l \leq t < t_k$ (Allison, 1995; Kaplan och Meier, 1958). Detta är samma sak som att för varje given tidpunkt t ta alla händelsetider som är mindre eller lika med t , räkna ut den betingade sannolikheten för att överleva till tidpunkt t_{j+1} givet att man överlevt till tidpunkt t och slutligen multiplicera dessa betingade sannolikheter tillsammans.

Kaplan-Meier kommer att användas för att skatta datamaterialets överlevnadsfunktion och den den logaritmerade kumulativa hazarden (som används för att testa proportionaliteten i Cox-modellen). För att skatta hazarden används istället *Life-table estimatorn* eftersom den är betydligt mer lätthanterlig för att skatta just hazarden, då Kaplan-Meier kan bli väldigt irreguljär och därför måste jämnas ut med passande estimator så att den blir tolkningsbar (Collett, 2003).

2.1.2 Life-table estimatorn av hazardfunktionen

För att kunna använda denna metod så måste tiden först grupperas i m olika intervall. När den genomsnittliga hazarden skattas i varje grupp för varje intervall tas antalet observerade händelser för det specifika intervallet och divideras med den genomsnittliga överlevnadstiden för intervallet. Låt antalet händelser för det j :te tidsintervallet vara lika med d_j . Sedan antas att n'_j är det genomsnittliga antalet individer utsatta för risk för en händelse i just det intervallet. Givet att hazarden är konstant för varje intervall kan den genomsnittliga överlevnadstiden räknas ut för intervallet genom att ta $(n'_j - d_j / 2)\tau_j$, där τ_j är längden på det j :te intervallet och d_j är antalet som upplever händelsen under intervallet. Life-table estimatorn för hazardfunktionen är då

$$h^*(t) = \frac{d_j}{(n'_j - d_j / 2) \tau_j}$$

för $t'_j < t < t'_{j+1}$ så att $h^*(t)$ blir en stegfunktion (Collett, 2003).

Både Kaplan-Meier och Life-table estimatorerna är välanvända icke parametriska metoder för att i huvudsak skatta individers överlevnad och risk för en specifik händelse. Problemet med Life-table estimatorm är dock antagandet att hazarden är konstant för varje intervall. Detta kräver att intervall skapas där detta antas vara uppfyllt, vilket är något godtyckligt. Det finns även icke-parametriska test för om olika grupper av individer med någon egenskap är signifikant skilda från varandra. Två sådana tester är *Log-rank-testet* och *Wilcoxon-testet*.

2.1.3 Log-rank och Wilcoxon

Testerna syftar till att kunna avgöra om det finns någon skillnad mellan olika gruppers överlevnadstid. Detta kan vara intressant för exempelvis en kategorisk förklarande variabel, där det finns minst två grupper som man vill kunna se skillnaden hos. Nollhypotesen i testen är att det inte finns någon skillnad mellan de olika gruppernas överlevnad. De båda teststatistikorna utgår från skillnaden mellan observerat och förväntat antal händelser. Log-rank statistikan beräknas enligt

$$\sum_{j=1}^r (d_{1j} - e_{1j}),$$

där d_{1j} är antal händelser i grupp ett och tidsperiod j , e_{1j} är det förväntade antalet händelser för grupp ett och tidsperiod j (Allison, 1995). Detta summeras över alla händelsetidpunkter som totalt är r stycken. Summeringen kvadreras sedan och divideras med den skattade variansen. Den resulterande teststatistikan är chi-två fördelad med antal grupper minus ett som frihetsgrader.

Det andra testet, Wilcoxon, har en statistika som beräknas enligt

$$\sum_{j=1}^r n_j (d_{1j} - e_{1j}),$$

där n_j är antalet individer som är utsatta för risk (möjligheten att drabbas av en händelse) vid en tidpunkt (Allison, 1995). Inkluderingen av n_j i formeln ovan gör att Wilcoxon testet blir viktat efter hur många som är utsatta för risk vid en given tidpunkt. Eftersom det är fler som utsätts för risk vid tidiga tidpunkter (då relativt få har drabbats av händelsen) betyder det att tidigare observationer får större vikt än senare med Wilcoxon testet. Skillnader mellan olika gruppers överlevnad som visar sig tidigt och därefter blir mindre tydliga med tiden, kommer därför vara mer signifikanta med Wilcoxon testet. Å andra sidan kommer grupper vars skillnad i överlevnad visar sig först sent i den observerade tidsperioden vara mer signifikanta med Log rank testet. Det behövs dock metoder som kan inkludera flera variabler med tillhörande koefficienter då en hel modell skattas för överlevnadstiden eller hazarden, en sådan metod är accelerated failure time modell.

2.2 Parametrisk modell

Accelerated failure time modellen (AFT) är en helt parametrisk modell som estimeras med maximum likelihood (givet att det finns censurering). Modellen antar att effekten av förklarande variabler är multiplikativa i förhållande till tiden det tar för en händelse att inträffa, och på det viset accelererar eller minskar överlevnadstiden. (Collett, 2003; Ghilagaber, 2005). Ofta liknas modellen vid en linjär regressionsmodell men med logaritmen av överlevnadstiden för den i :te individen i vänsterledet enligt följande,

$$\log T_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \lambda \varepsilon_i,$$

där ε_i är en slumpmässig avvikelse, och ($\beta_0 - \beta_k$ samt λ) är parametrar som ska estimeras. Överlevnadstiden logaritmeras för att undvika negativa värden och λ är en scale parameter som möjliggör att ε_i kan fixeras och λ istället står för variationen. Kovariaterna kan då förändra skalan på den underliggande distributionen utan att påverka dess läge. Det är också enligt denna modell som SAS programvara estimerar AFT modellen.

Att AFT är helt parametrisk innebär att en underliggande distribution för överlevnadstider måste anges. Detta är en skillnad från Cox proportionella hazard modell som gör att AFT får något säkrare skattningar med mindre standardfel (Allison, 1995). Det kan dock vara svårt att hitta en fördelning som passar bra till överlevnadstiderna. Det finns flera olika fördelningar av överlevnadstid som kan anges inom AFT modellen, de som diskuteras i denna uppsats är framförallt Extended generalized gamma (EGG), exponentiell, Weibull, Log-normal, Gamma och Log-logistic.

2.2.1 Fördelningar

2.2.1.1 Exponentiell

Den exponentiella distributionen är den enklaste av alla distributioner som kan anpassas i en parametrisk modell. Överlevnadstider som är exponentiellt fördelade kännetecknas framförallt av en konstant hazard enligt följande

$$h(t) = \lambda,$$

där λ är en konstant. Högre λ innebär därigenom högre hazard och omvänt en lägre överlevnad. Överlevnadsfunktionen är

$$S(t) = e^{-\lambda t}.$$

Ett högre värde av konstanten λ innebär att högerledet blir mindre och därigenom blir överlevnadstiden kortare.

2.2.1.2 Weibull

Weibull distributionen är en något mer avancerad distribution än den exponentiella. Den tillåter för såväl stigande som avtagande och konstant hazard. Detta beror på att det i en Weibull modell finns två parametrar som skall skattas, en shape parameter γ och en scale parameter λ . Det kan visas att när γ fixeras till 1 blir Weibull distributionen en exponentiell distribution. Den exponentiella distributionen är därför ett specialfall av Weibull distributionen. Överlevnadsfunktionen för Weibull distributionen är

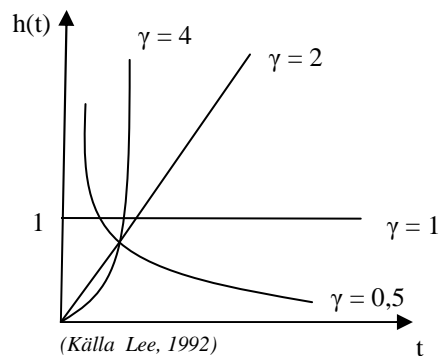
$$S(t) = e^{-(\lambda t)^\gamma}.$$

Hazard funktionen för Weibull distributionen är

$$h(t) = \lambda \gamma (\lambda t)^{\gamma-1}.$$

I båda dessa fall syns det tydligt att när shape parametern fixeras till 1 reduceras funktionerna till motsvarande överlevnads- och hazardfunktioner för den exponentiella distributionen. I figur 2.2 nedan visas funktionerna i grafisk form för olika värden av shape parametern.

Figur 2.2 Weibull hazard



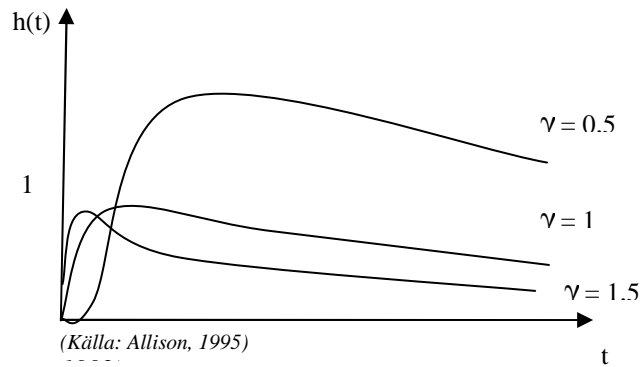
Specialfallet när γ är lika med 1 är även ett gränsvärde mellan stigande och avtagande risk i Weibull distributionen. När shape parametern γ överstiger 1 stiger hazarden i Weibull modellen medan en shape parameter under 1 innebär avtagande hazard.

2.2.1.3 Log-normal och Log-logistic

Om logaritmen av överlevnadstiderna är normalfördelade innebär det att överlevnadstiderna i sig själva har en Log-normal distribution. En skillnad från de tidigare distributionerna är att hazarden i Log-normal modellen är ickemonoton, med det menas att den kan vara både stigande och avtagande i samma funktion. Log-normal distributionen känns igen på dess inledningsvis kraftigt tilltagande hazard som sedan når en topp och därefter avtar. Detta syns i figur 2.3 nedan¹.

¹ Hazard- och överlevnadsfunktioner för Lognormal-, Loglogistic- och Gammamodellen finns i algebraisk form i bland annat Lee (1992), Allison (1995) och Collett (2003).

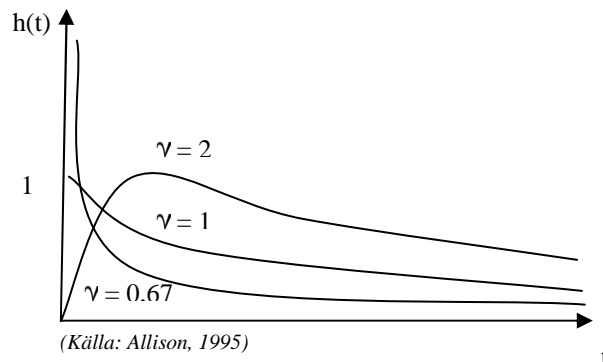
Figur 2.3 Log-normal hazard



Likt tidigare är det scaleparametern (λ) och shapeparametern (γ) som avgör utseendet på hazardkurvorna. Figur 2.3 ovan visar hur kurvorna skiljer sig med olika värden på shapeparametern (γ) när scaleparametern hålls konstant. Högre γ innebär att toppen inträffar tidigare och blir spetsigare.

Log-logistic fördelningen påminner om både Weibull och Log-normal modellerna. När dess shape parameter γ är under ett är hazardfunktionen monotont avtagande precis som Weibull. När γ istället är större än ett blir funktionen ickemonoton som för Log-normal fördelningen.

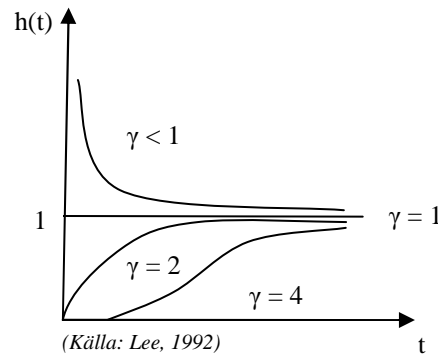
Figur 2.4 Log-logistic hazard



2.2.1.4 Gamma och EGG

Gamma distributionen har en shape parameter γ och scale parameter λ precis som flera av de tidigare fördelningarna. När λ fixeras vid exempelvis 1 som i figur 2.5 nedan kan man se hur hazarden varierar för olika värden av γ . En ytterligare likhet med Weibull distributionen är att när γ har värdet ett kommer risken att bli konstant. Det går alltså även här ta fram den exponentiella modellen som ett specialfall av Gammamodellen. När γ parametern är mindre än ett kommer hazarden att börja högt och sedan avta gradvis mot det värde som motsvaras av den exponentiella modellen (eller det värde som λ fixerats vid). Motsatsen gäller när γ parametern är större än ett, då är risken låg för tidiga observationer men stiger sedan för att också plana ut likt tidigare.

Figur 2.5 Gamma hazard



Extended generalized gamma (EGG) är en mycket mångsidig modell. Precis som den exponentiella modellen kan vara specialfall av Weibull och Gamma kan alla ovan nämnda modeller göras som specialfall av EGG (utom Log-logistic). Detta kan ske genom att shape (γ) och scale parametrar i EGG modellen fixeras vid olika värden beroende på vilken modell som ska efterliknas². Även standard normaldistribution kan sägas vara ett specialfall av EGG när shape parametern är fixerad till noll.

$$f(\gamma, \varepsilon) = \begin{cases} \frac{|\gamma|}{\Gamma(\gamma^{-2})} (\gamma^{-2})^{\gamma^{-2}} \exp\{\gamma^{-2}[\gamma\varepsilon - \exp(\gamma\varepsilon)]\}, & \gamma \neq 0 \\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2}\right), & \gamma = 0 \end{cases}$$

Nedan följer en lista med fixeringar av shape (γ) och scale (λ) parametrarna i EGG som reducerar EGG modellen till Exponentiell, Weibull, Gamma och Log-normal modellerna.

- $\gamma = 1$ och $\lambda = 1$ Exponentiell
- $\gamma = 1$ Weibull
- $\lambda = 1$ Gamma
- $\gamma = 0$ Log-normal

Fixeringarna som gjorts kan sedan hypotestestas för att kontrollera om fixeringarna (och därigenom även distributionerna) är rimliga. Testet är ett likelihood ratio test och förklaras i nästa stycke.

2.2.2 Goodness of fit test (likelihood-ratio)

Två olika distributioner kan testas mot varandra om den ena är ett specialfall av den andra. Själva testet utförs genom att beräkna den positiva skillnaden mellan de båda modellernas log-likelihood, skillnaden multipliceras sedan med två och det är detta som är teststatistikan. Teststatistikan är chi-två fördelad med antalet fixeringar som frihetsgrader. För att inte förlora en extra frihetsgrad när de dubbla fixeringarna görs vid testet av den exponentiella modellen, kan den exponentiella modellen istället testas som ett specialfall av Weibull där shape parametern är fixerad till ett. Nollhypotesen i testet är att de restriktioner som gjorts i modellen stämmer. En förkastad nollhypotes innebär alltså att de gjorda fixeringarna ej är signifikanta och inte kan sägas vara korrekta. Det innebär i sin tur att det är orimligt att den antagna distributionen ska gälla för de underliggande överlevnadstiderna.

² Scale parametern λ finns inte med i ovanstående EGG funktion utan tillförs genom AFT modellen som presenterades i avsnitt 2.2.

2.3 Semiparametrisk modell

En välanvänd icke parametrisk modell för att analysera överlevnadsdata är Cox proportionella hazard. Denna modell ger skattningar för den relativa risken att en händelse ska inträffa beroende på värdet av den förklarande variabeln. För att få fram denna modell så använder man något som kallas för *partial likelihood*, vilket inte är en sann maximum likelihood funktion eftersom den inte direkt använder sig av den verkliga överlevnadstiden för censurerade och icke censurerade fall (Collett, 2003). Denna funktion ser ut som följer med fixade parametrar från en proportionell hazard modell;

$$PL = \prod_{j=1}^r \left(\frac{\exp(\beta^T x_j)}{\sum_{k \in R_j} \exp(\beta^T x_k)} \right)^{\delta_j},$$

där R_j är risksetet för tidpunkt $t_j - 0$ (de individer som inte upplevt händelsen eller censureras till just den tidpunkten). Vidare är x_k en vektor med variabler för den k :te individen medan x_j är en vektor för de som upplever en händelse vid tidpunkt t_j och slutligen är β en vektor med koefficienter som ska skattas (Cox, 1975). Censurerade observationer exkluderas genom att sätta $\delta_j = 0$ för dessa (Collett, 2003). Denna formel är dock endast giltig då två händelser inte kan inträffa under samma tidpunkt. För att hantera detta problem med "tied data" (bunden data) finns det olika modeller som används beroende på om dessa *bindningar* kommer från att två händelser verkligen sker på samma tidpunkt eller om detta beror på att datamaterialet härstammar från grupperad ickebunden data. Datamaterialet i denna uppsats har en tidsvariabel som är grupperad efter år, trots att en händelse kan inträffa under olika tidpunkter under året. På grund av detta har en metod för bunden data vid namn *exact* använts, den utgår ifrån att händelserna sker under olika tillfällen trots att de är kodade likadant (Allison, 1995). Programvaran som används för att skatta Cox regressionsmodell kan även ge skattningar för samtliga variabelvärden, förutom det värde som används som jämförelsegrupp, för en kategorisk variabel³.

Om det nu antas att den linjära komponenten av en proportionell hazard modell innehåller p förklarandevariabler, X_1, X_2, \dots, X_p , och att de estimerade koefficienterna för dessa variabler är $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$. Om dessa logaritmeras fås den estimerade proportionella hazard modellen på följande vis

$$\log h_i(t) = \alpha(t) + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_p x_{ip}$$

där $\alpha(t)$ är lika med $\log \hat{h}_0(t)$ medan $\hat{h}_0(t)$ är den estimerade baseline hazard funktionen. För att kunna använda denna formel måste man först estimeras $\hat{h}_0(t)$ (Collett, 2003). Med Cox regression behövs inte någon distribution specificeras, eftersom maximering sker med partial likelihood, vilket gör att $\alpha(t)$ kan ta vilken form som helst (Allison, 1995). Detta gör Cox

³ Programmet i SAS heter Tphreg och är en utveckling av Phreg. Den har en klassificeringsprocedur som tillåter skattningar för olika variabelvärden, en procedur som även finns för motsvarande programvara för AFT-modeller.

regression relativt robust i jämförelse med accelerated failure time-modellen. Det kan dock vara en mindre lämplig modell om det finns starka skäl att tro att materialet härstammar från en specifik distribution. En begränsning är även antagandet att hazarden för en individ har en bestämd proportion gentemot alla andra individer. Om detta antagande inte är uppfyllt så faller modellen om proportionella hazards.

Det går dock att använda Cox regression utan antagandet om proportionella hazards. Detta kan man göra genom att specificera en tidsberoende variabel i modellen som kan anta olika proportioner över tiden. Ett annat sätt att hantera detta är genom att stratifiera efter en variabel som varierar med tiden samt är kategorisk och inte av direkt intresse⁴. Detta görs genom att konstruera separata partial likelihoodfunktioner för varje grupp som antar ett visst värde på stratifieringsvariabeln. Efter detta multipliceras dessa funktioner ihop och man väljer ett värde på β som maximerar funktionen (Allison, 1995). För att kontrollera om en variabel är proportionerlig eller inte krävs dock andra metoder.

2.3.1 Test av proportionaliteten i Cox-modellen

2.3.1.1 Den logaritmerade kumulativa hazarden

Den plottade logaritmerade kumulativa hazarden (LLS) kan användas för att kontrollera om en variabel är proportionerlig eller inte. Om en Cox regressionsmodell integreras över t och sedan logaritmeras ges följande ekvation

$$\log H_i(t) = \beta'x_i + \log H_0(t),$$

vilket säger att skillnaden i LLS-funktionen inte är beroende av tiden (Collett, 2003). Om LLS plottas för olika värden på en förklarande variabel så förväntas kurvorna vara proportionerliga över tiden. Om de inte är det så förkastas antagandet om proportionalitet. En annan aspekt är att LLS tar formen av en rak linje då materialet kommer från en Weibull distribution. Om den dock skulle vara konstant så är det ett tecken på att vårt material är exponentiell distribuerat (Collett, 2003).

Den logaritmerade kumulativa hazarden är lättförståelig både som modell och grafiskt, men dock till viss del subjektiv. Hur icke-proportionerlig ska olika LLS-kurvor vara för att modellen ska förkastas? Detta är inte alltid helt enkelt att svara på och därför kan andra metoder vara nödvändiga, även för att kunna hantera kontinuerliga förklarande variabler⁵.

2.3.1.2 Schoenfeld residualer

Schoenfeld residualer är inte beroende av den observerade överlevnadstiden vilket gör att man kan plotta den i :te residualen mot t_i för att testa hazardens proportionalitet. Det behövs inga estimat från hazardfunktionen vilket förenklar dess asymptotiska distribution (Schoenfeld, 1982). Man får heller inte bara en residual per individ utan en mängd residualer för varje förklarande variabel specificerad i en Cox regressionsmodell (Collett, 2003). Modellen ser ut som följer för den i :te residualen och den j :te förklarandevariabeln;

$$r_{Sji} = \delta_i |x_{ij} - \hat{a}_{ij}|,$$

⁴ Anledningen till att man väljer en variabel som inte är av direkt intresse beror på att det inte ges några skattningar för denna variabel då den används för stratifiering.

⁵ Det finns dock förslag på hur den logaritmerade kumulativa hazarden kan användas på ett mer precist sätt för att testa proportionalitet (Islam och Khondoker, 2009).

där x_{ij} är värdet på den j :te förklarande variabeln för den i :te individen i studien,

$$\hat{a}_{ij} = \left(\frac{\sum_{l \in R(t_i)} x_{jl} \exp(\hat{\beta}' x_l)}{\sum_{l \in R(t_i)} \exp(\hat{\beta}' x_l)} \right)$$

och där $R(t_i)$ är ett set med individer som är utsatta för risk under tid t_i . Parametern δ_i tar värdet noll om överlevnadstiden är censurerad och i annat fall värdet ett, vilket gör att Schoenfeld residualer endast är icke nollvärden om de kommer från ocensurerade observationer. Om däremot den största överlevnadstiden är ocensurerad så kommer den residualen att bli lika med noll (givet ekvationen ovan). För att skilja dessa från censurerade observationer så brukar de censurerade oftast uttryckas som saknade värden. Andra aspekter när Schoenfeld residualer skattas är att summan alltid blir noll och att förväntade r_{Sji} är noll. Residualerna är dessutom okorrelerade gentemot varandra i tillräckligt stora stickprov (Collett, 2003).

För att kunna använda Schoenfeld residualer för att testa en modell effektivt krävs det en slutlig justering. Låt vektorn för den i :te individen ta följande form $r_{Si} = (r_{S1}, r_{S2}, \dots, r_{Si})'$. Den vägda Schoenfeld residualer, r_{Si}^* , blir då komponenter av en vektor som följer

$$r_{Si}^* = r \text{ var}(\hat{\beta}) r_{Si},$$

där r är antalet händelser bland n individer och $\text{var}(\hat{\beta})$ är kovariansmatrisen med varianser för de estimerade parametrarna i en Cox regressionsmodell (Collett, 2003). Om $r_{Si}^* + \hat{\beta}_j$ plottas mot händelsetiden så fås information om formen på koefficienten för $\beta_j(t)$. En horisontell linje skulle då visa på att koefficienten är konstant, vilket gör att antagandet om proportionalitet inte kan förkastas. Även p-värden kan tas fram för att testa om det finns ett signifikant samband mellan residualerna (Grambsch och Therneau, 1994).

Problemet med detta test är att tolkningen kan bli olika beroende på om residualerna plottas mot händelsetiden eller en funktion av händelsetiden (exempel på detta ges senare då testet tillämpas). Övriga sätt att kontrollera om en Cox regressionsmodell är proportionerlig är att skapa tidsberoende variabler, och om de visar sig vara signifikanta så ger det oss en klar indikation om att proportionalitetsantagandet inte är uppfyllt. Då tecken visas på icke proportionerlighet kommer tidsberoende variabler att skapas för att kunna använda Cox regressionsmodell utan att anta proportionerlighet.

2.4 Multinomial logit-modell

I en logit-modell finns inga problem med "tied data" och tidsberoende variabler eftersom överlevnadstiden bryts ner till distinkta observationer. Det ges även möjlighet att testa hypoteser om det finns beroende över tid gällande hazarden. Modellen estimerar en binär regressionsmodell som säger om en händelse inträffade eller inte för varje tidsenhet. Denna modell kan användas både för diskret och kontinuerlig data (Allison, 1995).

Om P_{it} är en betingad sannolikhet att en individ i upplever en händelse under tidpunkten t , givet att ingen händelse har inträffat för individen, då får vi en modell för diskret data (kan

även användas när bindningar skapats då kontinuerlig tid grupperats) genom logistisk regression som ser ut som följer

$$\log\left(\frac{P_{it}}{1-P_{it}}\right) = \alpha_t + \beta'_k x_{itk},$$

där β' är en vektor med koefficienter och x_{it} är en vektor med variabler. Denna modell skattas genom maximum likelihood så att explicita estimat ges för α_t (Allison, 1995). Denna modell kan även utvecklas så att den gäller för flera möjliga händelser. Detta med antagandet om att händelserna inte är informativa gentemot varandra. Om detta antas ser multinomial logit-modellen ut på följande sätt

$$\log\left(\frac{P_{ijt}}{P_{i0t}}\right) = \beta'_j x_{it},$$

där P_{ijt} är den betingade sannolikheten att en händelse av typ j inträffar för individ i under tidpunkten t , givet att ingen händelse har inträffat för individen i fråga tidigare. Vidare är P_{i0t} sannolikheten att ingen händelse inträffar för individ i under tidpunkten t (Allison, 1995).

Fördelen med att använda denna modell är att det inte finns några problem med tidsberoende variabler och bunden data. Problemet är dock antagandet att händelserna inte får vara informativa gentemot varandra (vilket även gäller då det censureras i de tidigare modellerna), vilket kan vara fallet då det finns flera möjliga händelser. Ett annat problem med logit-modeller är att dataprocedurer inte har samma flexibilitet vad gäller val av distribution som procedurer för AFT.⁶ En viktig skillnad är att denna modell skattar betingade sannolikheten att en händelse ska inträffa en viss tidpunkt medan Cox regressionsmodell ger den relativa risken för en händelse och en AFT-modell skattningar på hur koefficienterna påverkar överlevnadstiden. Innan datamaterialet presenteras ges här några metoder för att bättre kunna hantera just konkurrerande risker.

2.5 Kumulativ incidens

För att få en beskrivande skattning av överlevnadstiden då hänsyn samtidigt tas till att det finns konkurrerande risker kan den kumulativa incidensen (CIF) användas som följer

$$\hat{F}_i(t) = \sum_{j, t_j \leq t} \hat{h}_{ij} \hat{S}(t_{j-1}),$$

där h_{ij} är den specifika hazarden för händelse i under tidpunkten t_j . Vidare är $\hat{S}(t_{j-1})$ en Kaplan-Meier estimator för sannolikheten att klara sig utan händelse före tidpunkt t_j . Givet att en individ inte upplevt någon händelse före tidpunkt t_{j-1} kan sannolikheten för en händelse i under intervallet t_{j-1} till t_j estimeras på följande vis

⁶ Det valda programmet för att analysera en multinomial logit-modell med två händelser är Probit i SAS, eftersom den har en klassificeringsprocedur som ger skattningar för flera olika variabelvärden för kategoriska variabler.

$$\hat{F}_i(t) = \sum_{j, t_j \leq t} \frac{d_{ij}}{n_j} \hat{S}(t_{j-1}),$$

där d_{ij} är antalet händelser av typ i som sker under tidpunkt t_j medan n_j är antalet som är utsatta för risk under samma tidpunkt. Detta gör att den kumulativa incidensen för en händelse av typ i inte bara beror på antalet individer som upplever den specifika händelsen utan även på antalet individer som inte upplevt någon annan typ av händelse. På så vis kan man uttrycka CIF som sannolikheten att en individ upplever en händelse i under tidpunkten t (Pintilie, 2006). Detta liknar den skattade överlevnadsfunktionen med Kaplan-Meier men här tas hänsyn till problematiken med konkurrerande risker. En andra skillnad är att CIF aldrig är avtagande, med andra ord är den kumulativ.

Vill man vidare inkorporera informationen för både händelsen av intresse och den konkurrerande händelsen så kan man använda sig av en betingad sannolikhet (CP). Den betingade sannolikheten beräknas för en tidpunkt t och är sannolikheten att en individ upplever en händelse av intresse under denna tidpunkt givet att individen inte upplever någon annan konkurrerande händelse. Om det vidare antas att det bara finns två olika typer av händelser – händelsen av intresse och den konkurrerande händelsen – så kan CP skattas för händelsen av intresse som

$$\hat{CP}(t) = \frac{\hat{F}_{ev}(t)}{1 - \hat{F}_{cr}(t)},$$

där \hat{F}_{ev} och \hat{F}_{cr} är de skattade CIF:en för händelsen av intresse och den konkurrerande händelsen. Detta gör att CP inte bara förändras när en händelse av intresse observeras utan även när en konkurrerande händelse inträffar (Pintilie, 2006).

2.5.1 Pepe och Moris metod

Pepe och Mori (1993) har utvecklat en metod där man kan jämföra kumulativa incidenskurvor för att testa effekten av en variabel. Testet är i huvudsak baserat på den vägda arean mellan de kumulativa incidenserna. Efter Pepe och Mori definierar Pintilie (2006) deras metod i det diskreta fallet för två överlevnadsfunktioner som

$$s = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \sum_{t_j} \{W(t_j) [\hat{F}_1(t_j) - \hat{F}_2(t_j)] (t_{j+1} - t_j)\},$$

där $F_i(t)$ är den kumulativa incidensfunktionen för händelsen av intresse för grupp i medan N_i är det totala antalet i grupp i . Den vägda funktionen $W(t)$ ges av

$$W(t_j) = \frac{(N_1 + N_2) \hat{C}_1(t_{j-1}) \hat{C}_2(t_{j-1})}{N_1 \hat{C}_1(t_{j-1}) + N_2 \hat{C}_2(t_{j-1})},$$

där $1 - \hat{C}_1(t)$ är Kaplan-Meier estimatorn för den censurerade distributionen. $C(t)$ är en överlevnadsfunktion där händelserna är definierade som observationer som antingen är censurerade eller en konkurrerad händelse. Den vägda funktionen minskar med tiden vilket gör att dess effekt på s blir mindre för skillnader i de kumulativa incidenserna längre fram i

tiden (Pintilie, 2006). Detta test har alltså en del likheter med Wilcoxon som presenterades tidigare med undantaget att Pepe och Moris metod kan användas då det finns konkurrerande risker. Detta test kan på grund av sin generella form utvecklas för att testa skillnader mellan två betingade sannolikheter genom att byta ut CIF:en mot CP:s (Pintilie, 2006).

3 Beskrivning av data

Datamaterialet som analyseras i denna uppsats, med metoderna som beskrivits i föregående avsnitt, är hämtade från en studie gjord av Bienen och van de Walle (1991). De har samlat in information om ledares tid vid makten över hela världen från början på 1800-talet till år 1987 där studien slutar. Sammanlagt är det 2257 observationer av ledare i 167 olika länder. Utöver detta finns även olika variabler som beskriver ledarna, men också variabler som är landsspecifika. I denna del ges först en beskrivning av de variabler som finns till förfogande och vad som bör beaktas vad gäller dessa. Därefter presenteras datamaterialet och problem som kan tänkas finnas och hur dessa i så fall kan hanteras. En presentation av variablerna och hur de är kodade finns i appendix A.

3.1 Variablerna

Det finns tre olika typer av händelser, antingen kan ledaren avlida av naturliga orsaker eller så kan han/hon förlora makten på konstitutionellt eller icke konstitutionellt vis. Ledarens tid vid makten är beräknat i år och den som inte har uppnått ett år vid makten har kodats som noll. Sedan finns några variabler som beskriver själva ledaren medan andra beskriver landet ledaren regerar över.

Saker som bör beaktas är att det inte finns information om årtal för variabelvärdena *islam*, *pop*, *literacy* och *ethnic*. Dessa siffror är konstanta över tiden då ett land har getts samma nivå på variablerna oavsett vilken tidpunkt ledaren regerade. Detta bör hållas i åtanke då slutsatser ska dras om hur exempelvis landets populationsstorlek eller läs- och skrivkunigheter påverkar risken att ledaren förlorar makten.

Vissa variabler gäller inte heller för hela datamaterialet utan endast för vissa regioner. Dessa är *colonial*, *order* och *islam*. Det går alltså inte att använda dessa för hela datamaterialet i uppsatsen. De ekonomiska variablerna (*income1*, *income2* och *growth*) är inte heller informativa för hela studieperioden utan bara slutet av mätperioden.

3.2 Ledarnas överlevnad för hela datamaterialet

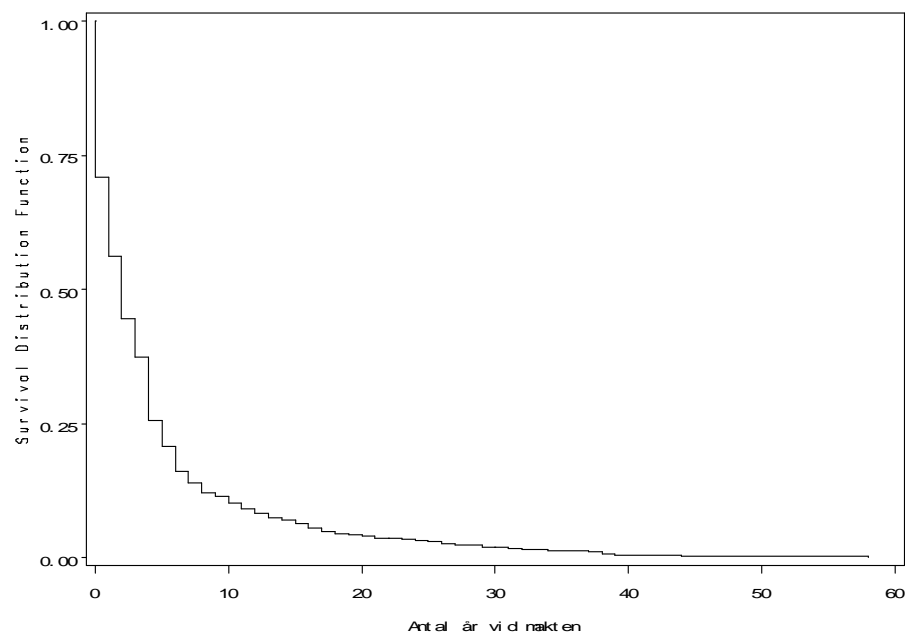
Av 2257 observationer så var det sammanlagt 2092 händelser av något slag medan 165 var censurerade. Vidare hade 132 förlorat makten genom att ha avlidit på naturlig väg medan 1394 förlorade makten på konstitutionell väg och 566 på icke konstitutionellt vis. Medellivslängden vid makten för alla ledare var 3,57 år och motsvarande median var 2,00 år. Många ledare förlorar alltså makten mycket tidigt i studien. Minsta överlevnadstiden för ledarna var 0 år, vilket 672 ledare hade (32 av dessa var censurerade), medan den längsta var 58 år⁷. Att det finns en tydlig snedfördrivning åt höger i datamaterialet kan redan utläsas av dessa siffror. Skattade överlevnadsfunktioner för samtliga ledare syns i figur 3.1. Life-table estimatorn för hazarden visas i figur 3.2 och med intervallen i tabell B.1 i appendix B.

⁷ Pedro II regerade i Brasilien från 1831 – 1889.

Figurerna och tabellen i appendix visar att cirka 90 procent av alla händelser sker innan det tionde året vid makten. Redan efter det fjärde året har 75 procent av alla ledare i studien förlorat makten.⁸

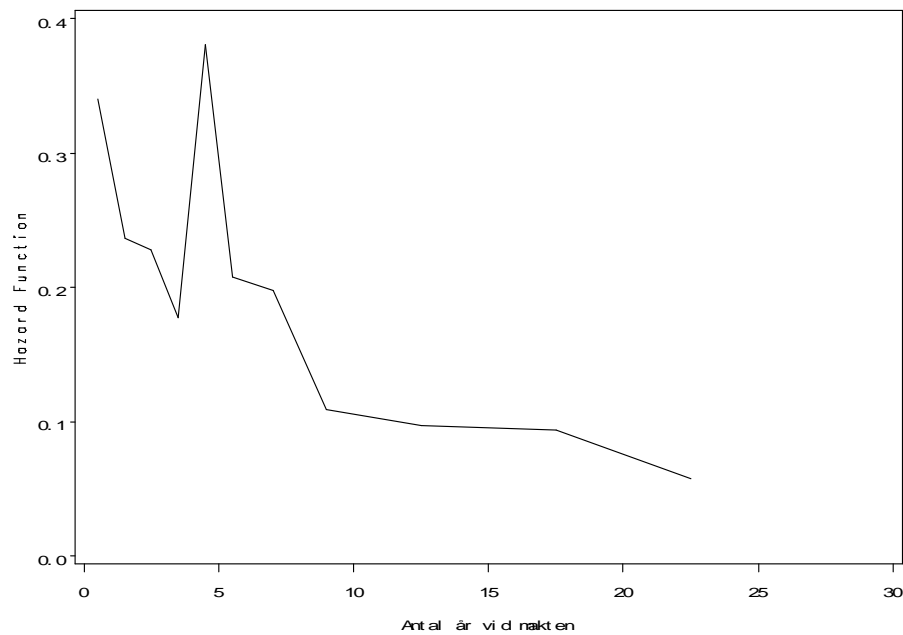
Som vi kan se så finns det en tydlig tendens till att hazarden för att förlora makten minskar med tiden, med undantag för år 4 då vi har en klar ökning. I figur B.1 i appendix har hazarden plottats för varje händelse genom att räkna de andra händelserna som censur. Detta tillvägagångssätt kan ifrågasättas eftersom censurering av vissa händelser inte tar hänsyn till att dessa även kan påverka utgången (en diskussion kommer senare i detta kapitel om informativ censurering). Hazarden verkar dock vara högst för konstitutionell utgång och lägst för naturlig död. Spiken under år fyra är främst på grund av konstitutionell utgång och delvis för icke konstitutionell. Hazarderna närmar sig varandra mot slutet och har en generellt avtagande trend, bortsett från naturlig död som är mer konstant och till och med ökar en del mot slutet (se vidare figur B.1 i appendix). Medelvärdet för antalet regerade år var 2,72 och medianen 2,00 för de som förlorade makten på konstitutionellt sätt. För de som förlorade makten på icke konstitutionellt vis var motsvarande siffror 3,33 och 1,00 år medan de som förlorade makten genom naturlig död hade medelvärde 7,48 och median 3,5 år.

Figur 3.1 Överlevnaden för samtliga ledare med Kaplan-Meier



⁸ Skattningarna av överlevnadstiden och hazarden är gjorda med proceduren Lifetest i SAS.

Figur 3.2 Hazarden för samtliga ledare med Life-table



Intressant är även att se hur makttillträdet och ledarens bakgrund (militär eller civil) kan påverka hur ledaren sedan har förlorat makten. En kort presentation av variablerna finns i tabell B.2 i appendix. I tabellen finns alla observerade händelser och censureringar uppdelade efter vilken typ av makttillträde ledaren gjorde (konstitutionell eller inte) samt om ledaren har militär eller civil bakgrund. Den största posten i tabellen är för civila ledare som tillträtt konstitutionellt och även blivit av med makten på konstitutionell väg. Strax under hälften av datamaterialet, 1030 stycken, finns i den kategorin. Två andra poster som kan vara värda att nämna är de civila (militära) ledare som fått makten via konstitutionella (icke konstitutionella) vägar och blivit av med den på ett icke konstitutionellt sätt. Vardera av dessa kategorier har några fler än 200 observationer vilket tillsammans är strax under 20 procent av hela datamaterialet. Alla tre posterna tar alltså upp nästan 70 procent av alla observationer.

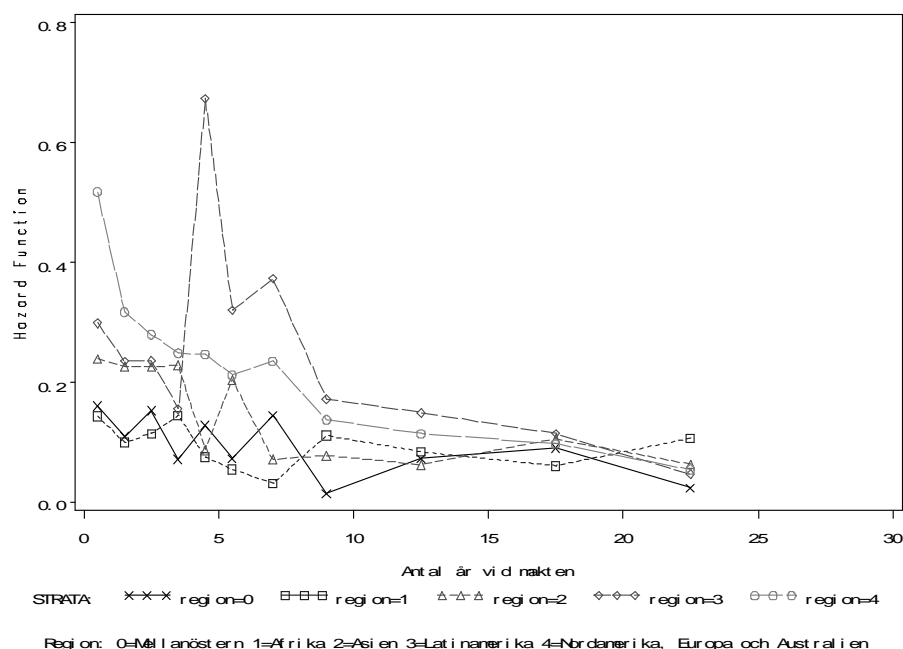
3.3 Ledarnas överlevnad i regionerna

I figur 3.3 har regionernas hazard plottats gentemot varandra. I appendix finns medelvärden och median för regionerna samt frekvensdistributionen för antalet år vid makten uppdelat mellan regionerna (tabell B.3 och B.4). Utifrån figuren syns att Latinamerika är den region med högst hazard från och med år 4 medan Nordamerika, Europa och Australien har den högsta hazarden fram till detta år. Överlevnadstiden är kortast i Nordamerika, Europa och Australien medan Afrika och Mellanöstern verkar ha de ledare som regerar längst (se tabell B.3). Från figur 3.3 (och tabell B.4) kan även utläsas att det är framförallt Latinamerika som står för spiken år fyra och därmed även spiken för hela materialet. Detta beror inte bara på att många förlorar makten detta år i Latinamerika utan även för att Latinamerika är en stor del av hela materialet (45 procent av observationerna) och därför även har stor påverkan när skattningar görs för hela populationen. För att visa detta grafiskt har även hazarden skattats för alla regioner utom Latinamerika i figur B.2 i appendix. I denna figur har spiken år 4 praktiskt taget försvunnit och det är en jämnare nedgång för hazarden över tiden.

Tabell B.5 i appendix visar hur de olika händelserna är fördelade över regionerna. Konstitutionell utgång är den vanligaste händelsen för alla regioner utom Afrika och

Mellanöstern, där är det vanligare att ledarna förlorar makten på icke konstitutionellt vis. Icke konstitutionell maktförlust är även ofta förekommande i Latinamerika där nästan en tredjedel av ledarna förlorat makten på detta sätt. I Nordamerika, Europa och Australien finns en tydlig dominans för konstitutionella maktförluster, övriga händelser är marginellt förekommande.

Figur 3.3 Hazarden för de olika regionerna med Life-table



3.4 Konkurrerande risker

I materialet finns tre möjliga händelser; konstitutionell maktförlust, icke konstitutionell maktförlust och slutligen naturlig död. Om endast konstitutionell maktförlust ska undersökas så kan de andra två händelserna (icke konstitutionell och naturlig död) ses som konkurrerande risker gentemot denna. Detta blir ett problem i exempelvis figur B.1 då hänsyn ej tas till de andra konkurrerande riskerna när hazarden skattas för varje händelse. Naturlig död skiljer sig betydligt från de andra två händelserna, dessutom är det få som förlorar makten på detta sätt. Att censurera naturlig död bör då inte vara något större problem om det även antas att denna händelse inte är informativ. Att göra likadant för de två andra händelserna och censurera dessa när man analyserar händelsen av intresse är dock betydligt svårare att försvara då de inte bara är lika varandra utan även kan ses som substitut till varandra. Risken för att dessa händelser är informativa gentemot varandra kan alltså orsaka bias i skattningarna om den ena händelsen censureras för att analysera den andra.

Metoder för att hantera konkurrerande risker presenterades i avsnitt 2.5. En deskriptiv metod som sakta har ersatt Kaplan-Meier estimatorn under senare år är den kumulativa incidensfunktionen när man har konkurrerande risker (Pintilie, 2006).⁹ I figur B.3 i appendix finns den kumulativa incidensen plottad för samtliga konkurrerande risker. Där syns det att sannolikheten är betydligt högre för att en ledare ska förlora makten på konstitutionellt vis än de andra händelserna. Sannolikheten för att förlora makten genom naturlig död är mycket låg. Liknande slutsats kan dras genom figur B.4 i appendix över de betingade sannolikheterna för

⁹ För att analysera kumulativa incidenser och den betingade sannolikheten för händelserna har *macron cuminc* använts i SAS, framtagen av Pintilie (2006).

varje händelse. Den betingade sannolikheten är högst för konstitutionell utgång och lägst för naturlig död. Dessa två figurer (figur B.3 och B.4) tillsammans med figur B.1 i appendix ger en sammanhängande bild av hur händelserna för ledarnas maktförlust förhåller sig till varandra.

3.5 Vänstercensurering

Som nämndes tidigare finns det många ledare (672 stycken) som vart vid makten i mindre än ett år när de förlorat den. Detta innebär att de i datamaterialet finns noterade med noll år vid makten, men i verkligheten kan de ha varit vid makten mellan noll och tolv månader. Denna dåliga precision skapar en del problem för vissa databehandlingsmetoder som används i uppsatsen. Som exempel används inte nollexponeringarna alls av SAS metoden Lifereg och resultaten blir då missvisande. Detta är särskilt viktigt för denna uppsats eftersom så stor del av materialet är observationer som ligger relativt tidigt. En lösning på problemet är att istället för noll tidsperioder ange 0,5 tidsperioder för de ledare som varit vid makten i mindre än ett år. Denna metod är dock något godtycklig och den ger lika stor vikt för en ledare som varit vid makten i en månad som tolv månader. En annan något mer komplicerad metod är att tillämpa vänstercensurering för observationer innan år ett (Allison, 1995). Detta görs när det inte riktigt går att säga när en händelse sker förutom att den sker innan en viss tidpunkt, i detta fall år ett. Vänstercensureringen utgår ifrån den befintliga överlevnadsfunktionen och dess likelihood, för alla tidpunkter efter ett, när den inkluderar de observationer som har nollexponering i modellen. Metoden innebär ändå en viss informationsförlust eftersom de exakta tidpunkterna inte går att ange, men informationsförlusten är mindre än utgångsläget när nollexponeringarna exkluderats helt.

3.6 Naturlig död

Som nämnts tidigare bör det inte vara något problem att censurera naturlig död om denna händelse kan ses som icke informativ gentemot de andra händelserna eftersom denna händelse är marginellt förekommande. Naturlig död bör hanteras annorlunda i jämförelse med de andra händelserna eftersom det kan antas att det är helt andra variabler som påverkar om man dör av naturliga orsaker, såsom ledarens ålder eller hälsa. Om händelsen av intresse är att en ledare blir avsatt av en konkurrent eller av någon/några andra så bör naturlig död hållas censurerad så länge det kan antas att naturlig död inte är informativ. Bienen och van de Walle (1991) själva har valt att i vissa fall hantera naturlig död på detta sätt. Även Allison (1995) gjorde detta i sin mindre studie av datamaterialet. Argumentet mot att censurera naturlig död är att en viss information går förlorad. Detta är dock inte ett tillräckligt starkt skäl för att inte censurera naturlig död eftersom skattningarna inte bara skulle bli mindre signifikanta utan även blir missvisande då naturlig död inte kan likställas med de två andra händelserna. Därför kommer naturlig död att censureras framöver när modeller ska skapas för datamaterialet.

4 Resultat

I detta avsnitt används metoderna som beskrivits i avsnitt 2 för att bearbeta datamaterialet. Lämpligheten kommer att testas för de olika modellerna, vilka distributioner som passar datamaterialet samt hur variablerna påverkar överlevnadstid och risken att förlora makten.

4.1 Icke parametriska test för variablerna

Till att börja med presenteras här några icke-parametriska test för de kategoriska variablerna som gäller för hela datamaterialet. Detta test är utfört för konstitutionell och icke-konstitutionell utgång tillsammans som händelse medan naturlig död censurerats. Resultaten av testen syns i tabell 4.1 med tillhörande frihetsgrader, chi-två och p-värden.

Tabell 4.1 Log-rank och Wilcoxon test över variablerna med tillhörande chi-två och p-värde i parantes

Variabel	Frihetsgrader	Log rank	Wilcoxon
Manner	1	1,34 (0,25)	13,29 (0,00)
Military	1	0,24 (0,63)	0,22 (0,64)
Ethnic	2	67,14 (0,00)	72,76 (0,00)
Region	4	183,94 (0,00)	155,58 (0,00)
Multiple	1	76,27 (0,00)	38,36 (0,00)

Variablerna *ethnic*, *region* och *multiple* är signifikanta för båda testen på 1 procentnivå. Detta innebär att det finns skillnader på överlevnadstiden för de olika nivåerna inom variablerna. *Military* är å andra sidan icke-signifikant för båda testerna. Det finns alltså ingen skillnad mellan ledarens militära eller civila bakgrund vad gäller tiden vid makten. *Manner* har icke-signifikant värde på Log-rank testet, men är däremot signifikant för Wilcoxon på 1 procentnivå. Detta beror på den tidigare diskussionen om skillnaderna mellan testen. För *manner* är skillnaden mellan variabelgrupperna större i början av undersökningsperioden vilket Wilcoxon testet lägger större vikt vid¹⁰.

Resultatet från tabell 4.1 gör som sagt ingen skillnad mellan händelserna konstitutionell och icke konstitutionell maktförlust. När händelserna behandlas separat blir resultaten något annorlunda, se tabell C.1 och C.2 i appendix. Alla variabler blir signifikanta på 1 procentnivå för både konstitutionell och icke konstitutionell maktförlust, utom variabeln *multiple* som endast blir signifikant för Wilcoxon på 5 procentnivå för icke-konstitutionell maktförlust.

4.2 Semiparametrisk modell

4.2.1 Test av proportionalitetsantagandet

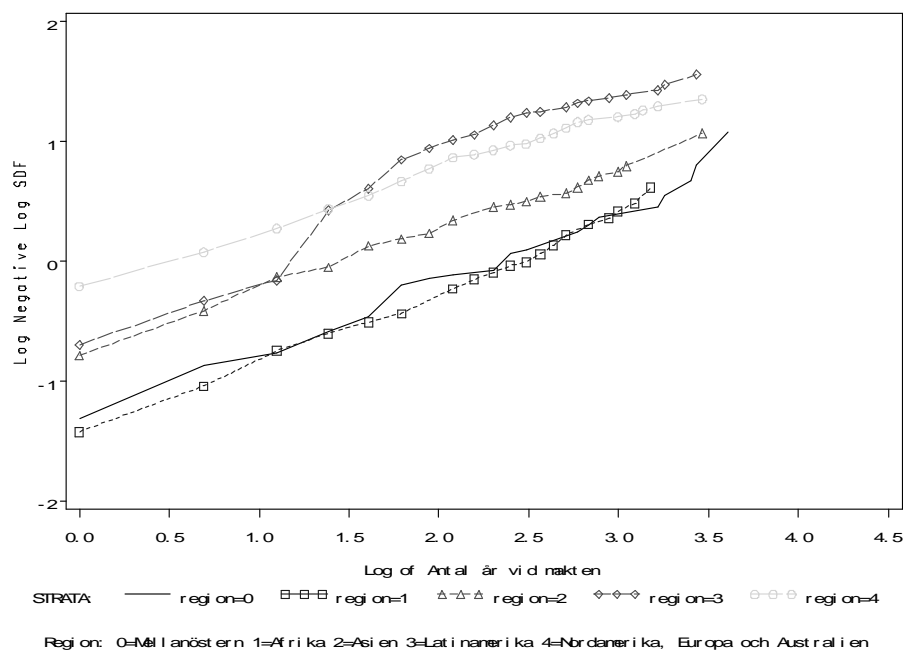
Två metoder för att testa antagandet om proportionalitet är den logaritmerade kumulativa hazarden (LLS) och Schoenfelds residualer, vilka har nämnts tidigare. För att få fram LLS har proceduren Lifetest använts med Kaplan-Meier skattningar. För att ta fram Schoenfelds

¹⁰ Samma test med naturlig död inkluderat resulterade endast i viss försvagning av signifikansen på variablerna, men inte tillräckligt för att någon nollhypotes som förkastades tidigare på 5 procentnivån skulle förkastas då naturlig död inkluderades. Detta bekräftar det tidigare diskuterade antagandet om att naturlig död skiljer sig från de andra två händelserna och därför inte tillför något till analysen av maktförlust i denna uppsats.

residualer har två macron använts i SAS¹¹. Denna procedur ger oss både p-värden och plottar för residualerna. Testen har utförts för konstitutionell och icke-konstitutionell maktförlust som gemensam händelse.

Då LLS plottas för alla olika värden på en kategorisk variabel förväntas skillnaden mellan grupperna vara konstanta över tiden när proportionalitet antas. Ett exempel när så inte är fallet syns i figur 4.1 för de olika regionerna. Här är Nordamerika, Europa och Australien (region 4) högst fram tills cirka år 4 då Latinamerika (region 3) korsar region 4 och håller sig högst perioden ut. Detta gör att proportionaliteten för denna variabel kan ifrågasättas.

Figur 4.1 Den logaritmerade kumulativa hazarden för variabeln *Region skattat* med Kaplan-Meier (naturlig död är censurerat).



En anpassning för Schoenfelds residualer visade på ett visst signifikant samband på 5 procentnivå då residualerna plottades mot $1 - \text{Overall Kaplan-Meier}$ ($1 - \text{KM}$), vilket är den förinställda transformationen av tiden i programvara som *R* (Pintilie, 2006). Dock är inte *region* signifikant på 5 procentnivå för transformationen rank av tiden (se figurer C.1 i appendix). Schoenfeld residual-test har en viss känslighet för vilken transformation som används för tidsaxeln vilket gör att slutsatser kan bli tveksamma när resultaten, som i detta fall, är olika vad gäller att förkasta eller inte förkasta proportionalitetsantagandet. För att vara på den säkra sidan kommer *region* användas som stratifieringsvariabel när Cox regressionsmodell anpassas för konstitutionell och icke-konstitutionell maktförlust tillsammans, med den konsekvensen att inga skattningar kommer att ges för denna variabel.

Logaritmen av den kumulativa hazarden för variabeln *multiple* visade på ett ökande avstånd mellan grupperna senare under tidsperioden. Detta betyder att risken ökar med tiden för gruppen som suttit i makten mer än en gång i jämförelse med de ledare som sitter sin första

¹¹ Macron som tar fram residualerna heter *schoen* och är framtagen av Bergstralh, Therneau och Lennon (2009). Denna macro använder sig i sin tur av macron *daspline* skapad av Harrell (1991). Testet har utförts med rekommenderat antal frihetsgrader och både med rankning av tiden (vilket är den förinställda transformationen) på x-axeln samt $1 - \text{Overall Kaplan-Meier}$.

gång vid makten. Schoenfelds residualer hade signifikans för både transformationen rank och 1 – KM (se tabell C.2 i appendix). När Cox modell skapas kommer en tidsberoende variabel skapas för *multiple* vilket även blir ett ytterliggare test för proportionalitetsantagandet.

För övriga variabler kunde inte antagandet om proportionalitet förkastas med LLS och Schoenfelds residualer. Schoenfelds residualer har även plottats för konstitutionell och icke konstitutionell maktförlust var för sig. Detta resulterade i att antagandet om proportionalitet förkastades för *multiple* för den tidigare händelsen, men inte för den senare. Variabeln *region* var icke-signifikant för båda dessa händelser. Givet detta kommer skattningar för regionerna ges för båda dessa händelser medan en tidsberoende variabel för *multiple* endast kommer skapas för konstitutionell maktförlust.

4.2.2 Anpassning av Cox regressionsmodell

En Cox regressionsmodell har skapats för konstitutionell och icke-konstitutionell maktförlust tillsammans. I tabell 4.2 syns variablerna och dess hazard ratio (riskkvot) med tillhörande chi-två och p-värden. Modellen är stratifierad efter regionerna och en tidsberoende variabel har skapats genom att ta variabeln *multiple* multiplicerat med tiden. Denna variabel heter *time-multiple* och är signifikant på 1 procentnivå vilket stärker de tidigare aningarna om att variabeln inte är proportionell. När Cox regressionsmodell anpassats har sista variabelvärdet för de kategoriska variablerna använts som baseline när skattningarna gjorts. Endast två variabler är icke signifikanta på 5 procentnivå i modellen, dessa är *military* och *literacy*. Variablerna *age*, *population* och *land* är signifikanta på 5 procentnivå, men riskkvoterna är så gott som obetydliga.

Om en ledare har tagit makten på konstitutionell väg är risken att förlora makten 20 procent mindre än om maktövertagandet hade varit på icke konstitutionellt vis. Risken är dock högre för ledare som regerar i ett land med mellan eller låg etnisk konflikt i jämförelse med ledare som regerar över länder med hög etnisk konflikt. Ledare som regerar för första gången har cirka 15 procent lägre risk att bli av med makten än ledare som regerat minst två gånger.

Tabell 4.2 Cox regressionsmodell för konstitutionell och icke-konstitutionell maktförlust gemensamt (stratifiering efter region).

Variabel	Kategori	Hazard Ratio	Chi-två	P-värde
Manner	Konstitutionell	0,80	12,48	0,00
Military	Civil	1,04	0,31	0,58
Ethnic	Låg	1,26	6,82	0,01
Ethnic	Mellan	1,29	8,10	0,00
Multiple	Första gången	0,84	6,90	0,01
Age		1,02	49,16	0,00
Literacy		1,00	1,06	0,30
Population		1,00	5,03	0,03
Land		1,00	18,75	0,00
Time-multiple		1,05	9,34	0,00

När Cox regressionsmodell anpassas för de separata händelserna konstitutionell och icke-konstitutionell utgång, blir resultatet delvis annorlunda. Nu är variabeln *region* inkluderat i båda modellerna eftersom proportionalitetsantagandet inte kunnat förkastas. Eftersom inget tidsberoende kunnat påvisats för variabeln *multiple* för icke konstitutionell maktförlust har inte heller variabeln *time-multiple* inkluderats i modellen (ett försök att inkludera denna variabel resulterade i ett p-värde som var långt ifrån signifikant). Modeller för vardera händelser syns i tabell 4.3.

Figur 4.3 Cox regressionsmodell för separerade händelser

Variabel	Kategori	Konstitutionell		Icke-Konstitutionell	
		Hazard Ratio	Chi-två (<i>p</i> -värde)	Hazard Ratio	Chi-två (<i>p</i> -värde)
Manner	Konstitutionell	1,34	10,46 (0,00)	0,42	74,85 (0,00)
Military	Civil	1,11	1,67 (0,20)	0,94	0,38 (0,54)
Ethnic	Låg	1,39	7,58 (0,01)	0,95	0,16 (0,69)
Ethnic	Mellan	1,30	4,61 (0,03)	1,15	1,21 (0,27)
Region	Mellanöstern	0,34	30,32 (0,00)	1,90	6,61 (0,01)
Region	Afrika	0,20	38,06 (0,00)	2,26	11,46 (0,00)
Region	Asien	0,61	18,75 (0,00)	1,58	3,62 (0,069)
Region	Latinamerika	0,78	11,06 (0,00)	3,62	51,45 (0,00)
Multiple	Första gången	0,76	13,54 (0,00)	0,96	0,17 (0,68)
Age		1,02	49,48 (0,00)	1,00	1,78 (0,18)
Literacy		1,01	16,53 (0,00)	0,99	10,12 (0,00)
Population		1,00	2,25 (0,13)	1,00	4,36 (0,04)
Land		1,00	16,81 (0,00)	1,00	2,84 (0,09)
Time-multiple		1,04	5,39 (0,02)	-	-

För modellen med risken att förlora makten på konstitutionell väg är nästan alla variabler signifikanta, dock ej *military* (ledarens bakgrund) och *population*. Trots signifikans är det även några variabler som har riskkvoter som ligger nära ett och därför kan sägas ha begränsad betydelse. Detta gäller variablerna *literacy* och *land*. De tydligaste effekterna finns hos variablerna *manner*, *ethnic* och *region*. En ledare som tagit makten konstitutionellt har 34 procent större risk att även förlora makten konstitutionellt än en ledare som tagit makten icke konstitutionellt. En låg eller mellan nivå av etnisk konflikt verkar öka risken för konstitutionell maktförlust med 30-40 procent i jämförelse med en hög grad av etnisk konflikt. Vad gäller *region* har alla andra regioner lägre risk för konstitutionell maktförlust än region 4 (Nordamerika, Europa och Australien). Mest markant är skillnaden för Afrika och Mellanöstern där risken är 80 respektive 66 procent lägre för en konstitutionell maktförlust än i region 4.

När risken för icke konstitutionell maktförlust modelleras finns både likheter och skillnader i resultatet. Fler variabler är inte signifikanta på 5 procentnivån, detta gäller *military*, *ethnic*, *region* (Asien), *multiple*, *age* och *land*. Likt tidigare har *literacy* och *population* riskkvoter nära ett. I denna modell finns de mest anmärkningsvärda effekterna hos variablerna *manner* och *region*. En ledare som kommit till makten konstitutionellt har 58 procent mindre risk att förlora makten icke konstitutionellt i förhållande till ett icke konstitutionellt makttillträde. Samtliga regioner har högre risk för icke-konstitutionell maktförlust än region 4. I Latinamerika och Mellanöstern är risken 262 respektive 126 procent högre i förhållande till region 4.

4.3 Parametrisk modell

I parametriska modeller kan särskilda distributioner för den underliggande överlevnadstiden anges. För att hitta en distribution/modell som passar till datamaterialet inleds avsnittet med distributionstester. Därefter skattas och tolkas den parametriska AFT modellen.

4.3.1 Distributionstest

I avsnitt 2.2 presenterades flera olika distributioner som kunde formuleras som specialfall av extended generalized gamma modell (EGG). Som beskrevs kan sedan modellerna testas med

ett log-likelihood test. Detta för att se om fixeringarna av parametrarna stämmer och det därigenom finns skäl att använda den underliggande distributionen av överlevnadstider som modell.

I tabell 4.5 nedan presenteras resultatet av ett sådant test när ingen skillnad gjorts mellan händelserna konstitutionell och icke konstitutionell maktförlust.

Tabell 4.5 *Log-likelihood test för fyra olika distributioner*

Distribution	Log-likelihood	Chi- två
EGG	-3122,69	-
Log-Normal	-3126,47	7,56
Weibull	-3178,5	111,62
Gamma	-3137,2	29,02
Exponential	-3349,29	341,58
(Log-logistic)	(-3132,08)	-

För att fixeringarna ska vara korrekta krävs att nollhypotesen inte förkastas i log-likelihood testet. Den skattade scale parametern för EGG var 1,29 och shape parametern var -0,37. Alla chi-två värden i tabell 4.5 är signifikanta på en procent nivå med en frihetsgrad. Det innebär att inga av de gjorda fixeringarna av scale och shape parametrarna kan sägas vara tillräckligt nära för att vara korrekta. Att Log-normal är den distribution som har minsta chi-två värde förklaras av att den fritt estimerade shape parametern i EGG är nära noll. En figur med EGG modellen anpassad till hazardkurvan för datamaterialet finns i appendix (figur C.3), där finns också motsvarande figur men med en Log-logistic modell anpassad (figur C.4)¹². Utifrån figurerna är det mycket svårt att bedöma om Log-logistic passar datamaterialet bättre än EGG, därför kommer två AFT modeller skattas.

I appendix C (tabell C.3 och C.4) finns även separata log-likelihood test för de två händelserna konstitutionell och icke konstitutionell maktförlust. Samtliga distributioner för båda händelser är signifikanta utom en, Log-normal modellen för icke konstitutionell maktförlust. Shape parametern för EGG är -0,23 och scale parametern 2,15 för icke konstitutionell maktförlust. Redan här kan ses att shape parametern är nära noll vilket talar för en Log-normal modell. Log-normal modellen hade i sin tur en scale parameter på 2,06 och ett chi-två värde mindre än 1 vilket inte tillåter att nollhypotesen förkastas. När överlevnadstiden för ledare ska modelleras och händelsen är icke konstitutionell maktförlust, kan en log-normal modell anpassas för att spara frihetsgrader.

Det finns även figurer i appendix (figur C.5 och C.6) där en EGG modell anpassats för risken för konstitutionell maktförlust och en Log-normal modell anpassats för risken för icke konstitutionell maktförlust. Log-logistic distributionen visade sig inte passa lika bra som EGG modellen för konstitutionell maktförlust. För icke-konstitutionell maktförlust var det återigen svårt att bedöma om Log-logistic passar bättre än resultatet från log-likelihood testet (Log-normal) och därför kommer både en Log-normal och Log-logistic modell anpassas.

4.3.2 Anpassning av AFT modell

Som tidigare har tre olika modeller anpassats, en för de gemensamma händelserna konstitutionell och icke konstitutionell maktförlust tillsammans (tabell 4.6) därefter två modeller där de två händelserna utreds var för sig (tabell C.5 och C.6)

¹² Figurer över distributionerna har gjorts med macron Lifehaz i SAS skapad av Allison (1995).

I tabell 4.6 finns resultaten från den första modellen där de två händelserna behandlas gemensamt. EGG användes som underliggande distribution efter svårigheten att bedöma Log-logistic modellen. Även en Log-logistic modell anpassades men gav inga väsentliga skillnader från EGG modellen i tabell 4.6. Som nämnts modellerar AFT *logaritmen av överlevnadstiden* i vänsterledet (till skillnad från Cox modellen där hazarden är i vänsterledet), estimaten i tabellen är därför koefficienternas påverkan på den logaritmerade överlevnadstiden.

Tabell 4.6 Resultat från AFT modell med båda händelser i en EGG modell.

Variabler	Estimat (β)	Standardfel	e^β	Chi-två	P-värde
Intercept	1,316	0,264		24,79	0,00
Manner (konstitutionell)	0,437	0,083	1,549	27,75	0,00
Military (civil)	-0,004	0,078	0,996	0,00	0,95
Ethnic 0 (låg)	-0,253	0,109	0,777	5,36	0,02
Ethnic 1 (mellan)	-0,233	0,107	0,792	4,76	0,03
Region 0 (Mellanöstern)	1,192	0,162	3,294	54,02	0,00
Region 1 (Afrika)	1,076	0,167	2,932	41,31	0,00
Region 2 (Asien)	0,773	0,122	2,166	40,44	0,00
Region 3 (Latinamerika)	0,313	0,084	1,367	13,90	0,00
Multiple (första)	0,307	0,067	1,360	21,09	0,00
Age	-0,021	0,003	0,979	52,71	0,00
Literacy	-0,002	0,002	0,999	0,61	0,44
Population	-0,001	0,000	0,999	6,82	0,01
Land	0,000	0,000	1,000	23,89	0,00
Scale	1,255	0,029			
Shape	0,093	0,118			

Mer lättolkad är kolumnen, e^β , där påverkan på *överlevnadstiden* är i förhållande till baselinekategorin när alla andra kovariater hålls konstanta¹³. Precis som tidigare är baselinekategorin den sista kategorin i variablerna. *Military* och *literacy* är inte signifikanta på fem procentnivån. Likaså är alla kontinuerliga variabler mycket nära ett och kan då sägas ha begränsad effekt på överlevnadstiden. En ledare som tagit makten på konstitutionell väg har cirka 55 procent längre överlevnadstid än en som tagit makten på icke konstitutionellt sätt. En låg eller mellan nivå av etnisk konflikt verkar minska tiden vid makten med cirka 20 procent jämfört med en hög nivå. Alla regioner har i modellen längre överlevnadstider vid makten än region 4. Mest markant är effekterna för Mellanöstern (cirka 230 procent längre), Afrika (cirka 190 procent längre) och Asien (cirka 115 procent längre). En ledare som är vid makten för första gången har ungefär 35 procent längre överlevnadstid jämfört med en ledare som vart vid makten minst två gånger.

I appendix (tabell C.5 och C.6) finns även tabeller för resultaten när varsin AFT modell skattades för de båda händelserna konstitutionell och icke konstitutionell maktförlust. För modellen med den konstitutionella händelsen har även en EGG modell anpassats. De största skillnaderna är att *manner* och *ethnic (mellan)* inte är signifikanta och effekterna från regionvariabeln går i samma riktning men är nu mycket större. Mellanöstern och Afrika har cirka 310 och 630 procent längre överlevnadstider än region 4 för konstitutionell maktförlust.

För modellen med icke konstitutionella maktförluster har en Log-normal modell anpassats, Log-logistic modellen visade sig inte ge några betydande skillnader precis som tidigare.

¹³ Kontinuerliga variabler som *Age*, *Literacy*, *Population* och *Land* tolkas som procentuell påverkan när den aktuella variabeln ökar med en enhet.

Skillnaden mot de tidigare modellerna är nu att *ethnic* (mellan och låg), *region* (Mellanöstern, Afrika och Asien) samt att *multiple* är icke signifikanta på fem procentnivån. Konstitutionellt makttillträde innebär cirka 275 procent längre överlevnadstider än icke konstitutionellt när alla andra kovariater hålls konstanta. Latinamerika har nu cirka 70 procent kortare överlevnadstider än region 4.

4.4 Multinomial logit modell

För att testa om det finns en signifikant skillnad för hur variablerna påverkar utgången mellan konstitutionell och icke konstitutionell maktförlust, givet att dessa inte är informativa, så har en multinomial logit-modell anpassats. Resultatet syns i tabell 4.7. Dessa koefficienter kan tolkas som skillnaden mellan koefficienterna för konstitutionell och icke konstitutionell maktförlust.

Tabell 4.7 Konstitutionell kontra icke konstitutionell maktförlust

Variabel	Estimat (β)	Standardfel	e^{β}	Chi-två	P-värde
Intercept	-1,71	0,52		10,68	0,00
Manner (konstitutionell)	1,13	0,15	3,08	59,45	0,00
Military (civil)	0,23	0,15	1,26	2,58	0,11
Ethnic (låg)	-0,17	0,23	0,84	0,56	0,45
Ethnic (mellan)	-0,18	0,22	0,83	0,69	0,41
Mellanöstern	-1,78	0,35	0,17	26,12	0,00
Afrika	-2,43	0,38	0,09	40,52	0,00
Asien	-1,08	0,29	0,34	13,98	0,00
Latinamerika	-1,38	0,21	0,25	45,23	0,00
Multiple	-0,23	0,14	0,79	2,62	0,11
Age	0,02	0,01	1,02	7,73	0,01
Literacy	0,03	0,00	1,03	45,5	0,00
Population	0,00	0,00	1,00	1,54	0,21
Land	0,00	0,00	1,00	0,01	0,92
Log-years ¹⁴	0,18	0,06	1,20	8,51	0,00

Likt AFT modellerna tidigare måste exponenterna av estimaterna beräknas för att bli mer tolkningsbara. Kolumnen med e^{β} visar således hur den betingade sannolikhetskvoten (odds) av två händelser (konstitutionell eller icke konstitutionell) påverkas av en variabel. För variabeln *manner* gäller då, givet att en ledare har tagit makten konstitutionellt och nu har förlorat makten, så är det ungefär tre gånger större sannolikhet att ledaren förlorat makten konstitutionellt än icke konstitutionellt. Variablerna *military*, *ethnic*, *multiple*, *population* och *land* är inte signifikanta på 5 procentnivån vilket betyder att det inte är någon signifikant skillnad på hur variablerna påverkar typen av maktförlust. Samtliga regionvariabler är signifikanta och visar på att sannolikheten är högre för icke konstitutionell maktförlust i förhållande till region 4. I jämförelse med region 4 har exempelvis en ledare från Afrika cirka tio gånger mindre sannolikhet att förlora makten konstitutionellt än icke konstitutionellt. För kontinuerliga variabler som *literacy* och *age* tolkas skattningarna som procentuell påverkan på sannolikhetskvoten. Ju äldre ledaren är desto större är oddsen för konstitutionell maktförlust, varje extra år ökar oddsen med två procent. En procent ökning av läskunnigheten ger en tre procentig ökning av oddsen för konstitutionell maktförlust.

¹⁴ Variabeln *log years* är en logaritmering av tiden som testar om scale parametern är lika för de två händelsernas distributioner (Allison, 1995).

4.5 Kumulativ incidens

Tidigare då varje händelse analyserats var för sig har övriga händelser helt censurerats. Konsekvenserna av detta beror på om man kan anta att händelserna är informativa gentemot varandra eller inte. Är de informativa så uppfylls inte ett av de grundläggande kraven för censurering. För att undvika detta och ändå kunna jämföra hur en variabel påverkar överlevnaden för en specifik händelse, samtidigt som hänsyn tas till att det finns en konkurrerande risk, presenterades en metod av Pepe och Mori i avsnitt 2.5¹⁵. Nedanför presenteras tabell 4.8 över resultaten av testen mellan CP-kurvorna. Variabeln *conflict* är variabeln *ethnic* då mellan och hög konfliktnivå slagits ihop och på så vis blivit en dikotom variabel. Detta på grund av att macron skapad av Pinitilie (2006) endast klarar av att genomföra testet för variabler som antar två olika värden.

Tabell 4.8 Pepe och Moris test för dikotoma variabler för konstitutionell samt för icke-konstitutionell maktförlust.

Variabel	Konstitutionell		Icke Konstitutionell	
	Chi-två	P-värde	Chi-två	P-värde
Manner	33,01	0,00	106,79	0,00
Military	19,25	0,00	66,58	0,00
Multiple	119,57	0,00	10,88	0,00
Conflict	66,02	0,00	2,43	0,12

Alla dessa fyra variabler visar sig vara signifikanta då hänsyn tas till den konkurrerande risken. Det är endast *conflict* för icke-konstitutionell utgång som inte är signifikant på 5 procentnivå. I övrigt skiljer sig inte resultatet i tabell 4.8 från de tidigare testen med Wilcoxon. I figur C.7 och C.8 i appendix presenteras CIF- och CP-kurvorna för variabeln *conflict* för icke-konstitutionell maktförlust då konstitutionell maktförlust behandlats som konkurrerande risk. CIF-kurvorna visar på samma signifikanta samband som för hazard-kurvorna för variabeln *ethnic*, lägre risk för låg etnisk konflikt, men när den betingade sannolikheten plottas med hänsyn till när den konkurrerande händelsen förändras blir resultatet icke signifikant. I figur C.8 är sambandet nästan det motsatta dessutom med högre risk för låg konfliktnivå. Detta ger en fingervisning om att det kan vara felaktigt att anta att händelserna är icke informativa gentemot varandra med avseende på variabeln *ethnic/conflict*.

En sammanfattning över resultaten från de olika Cox och AFT modellerna presenteras i tabell 4.9 nedan.

¹⁵ För att analysera detta har tre macron av Pinitilie (2006) används. Dessa ger plottar för både CIF och CP, men även test för hur kurvorna skiljer sig från varandra för en dikotom variabel. Detta test är baserat på Pepe och Moris metod (1993).

Tabell 4.9 Sammanfattning av modellresultat

Variabel	Hela materialet		Konstitutionell		Icke konstitutionell	
	Hazard ratio	AFT e^{β} (EGG)	Hazard ratio	AFT e^{β} (EGG)	Hazard ratio	AFT e^{β} (Log-normal)
Manner (konstitutionell)	0,80**	1,55**	1,34**	0,91	0,42**	3,75**
Military (civil)	1,04	0,99	1,11	0,93	0,94	1,09
Ethnic 0 (låg)	1,26*	0,78*	1,39*	0,70*	0,95	0,98
Ethnic 1 (mellan)	1,29**	0,79*	1,30*	0,78	1,15	0,78
Region 0 (Mellanöstern)	-	3,29**	0,34**	4,11**	1,90*	0,79
Region 1 (Afrika)	-	2,93**	0,20**	7,28**	2,26**	0,58
Region 2 (Asien)	-	2,17**	0,61**	2,05**	1,58	0,99
Region 3 (Latinamerika)	-	1,37**	0,78**	1,52**	3,62**	0,30**
Multiple (första)	0,84*	1,36**	0,76**	1,48**	0,96	1,11
Age	1,02**	0,98**	1,02**	0,98**	1,00	0,99*
Literacy	1,00	1,00	1,01**	0,99**	0,99**	1,01**
Population	1,00*	1,00*	1,00	1,00	1,00*	1,00*
Land	1,00**	1,00**	1,00**	1,00**	1,00	1,00*
Time-multiple	1,05**	-	1,04*	-	-	-

* = Signifikant på 5 procentnivå. ** = Signifikant på 1 procentnivå.

5 Analys och diskussion

Analysen är indelad i tre delar, först diskuteras modellernas variabler. Därefter vilken betydelse de konkurrerande händelserna och informativ censurering har på datamaterialet. Sist diskuteras vilken modell som verkar vara lämpligast för att analysera ledares tid vid makten.

5.1 Variablerna

Variabeln *manner* är signifikant för alla modeller utom AFT modellen för konstitutionell maktförlust. Bienen och van de Walles (1991) hypotes var att ledare som tagit makten konstitutionellt skulle ha högre legitimitet och på så vis mindre risk. Resultatet visar dock på att detta är tveksamt. Hur ledaren har tagit makten verkar ha mer betydelse för risken och överlevnadstiden vid icke konstitutionell maktförlust. Resultaten för hela materialet hamnar ofta någonstans mellan resultaten för de separata händelserna. När effekterna är starkare för *manner* vid icke konstitutionell maktförlust återspeglas även detta för hela materialet. Den lägre risken för de båda maktförlusterna för konstitutionella ledare kan bero på att effekten är betydligt starkare för icke konstitutionell maktförlust. Överlag pekar resultaten på att ledare som tagit makten konstitutionellt också kommer förlora makten konstitutionellt, omvänt är det också troligare att en ledare som tagit makten icke konstitutionellt också förlorar makten på det sättet. Om detta beror på legitimitet eller inte är svårt att svara på utifrån resultaten.

Ledare som har civil bakgrund har större risk (kortare överlevnadstid) för konstitutionell maktförlust och tvärtom för icke konstitutionell maktförlust. Variabeln är däremot inte signifikant för någon av modellerna vilket är något oväntat. Ledarens militära/civila bakgrund visade sig dock vara signifikant med Log-rank och Wilcoxon testen. Även en Cox regressionsmodell för vardera händelsen för sig visade sig vara signifikant för ledarens bakgrund, men endast då regionerna inte inkluderades. I appendix D finns separata modeller för regionerna (tabell D.1 till D.4) och visar endast på signifikant resultat för ledarens bakgrund vad gäller konstitutionell maktförlust i region 4 (tabell D.4). Risken är 90 procent högre för civila ledare att förlora makten konstitutionellt i jämförelse med militära ledare.

Sambandet är liknande för övriga regioner, men ej signifikant. I tabell D.5 i appendix har ledare som regerat från 1960 och framåt modellerats enskilt. Där visar sig variabeln *military* vara signifikant för icke konstitutionell maktförlust, men sambandet är det motsatta från tidigare. Risker är 60 procent större att förlora makten icke konstitutionellt som civil ledare. Bienen och van de Walle (1991) ville undersöka om det fanns någon skillnad vad gäller risk för ledare med dessa bakgrunder. Denna uppsats visar på bristande signifikans och motsäggande resultat för att några slutsatser kan kunna dras för hela materialet.

Ethnic verkar ha störst betydelse för de konstitutionella modellerna och visar att ledare som regerar i länder med låg eller mellan nivå av etnisk konflikt har högre risk (kortare överlevnadstid) i jämförelse med en hög nivå. Detta är något förvånande eftersom antagandet är att det är svårare att regera länder med en hög nivå av etnisk konflikt och därför borde risken vara mindre (eller längre överlevnadstider) i länder med låg nivå av etnisk konflikt. I analysen av regionerna (tabell D.1 till D.4) visar resultaten att Latinamerika (och till viss del även region 4) har starka effekter för konstitutionell maktförlust som går i samma riktning som resultatet ovan i tabell 4.9. Mellanöstern och Afrika har däremot motsatt samband, lägre risk för maktförlust vid mellan eller låg nivå av etnisk konflikt. Hur graden av etnisk konflikt påverkar ledares tid vid makten beror alltså i sin tur även på vilken region ledaren befinner sig i. Som tidigare påpekats är det dock långt ifrån jämnt fördelade observationer över regionerna. Den höga risken för ledare att förlora makten konstitutionellt i länder med låg och mellan etnisk konflikt, skulle kunna förklaras av att den lite etnisk konflikt troligtvis ökar chansen för ett demokratiskt statskick. Därigenom blir också risken att förlora makten konstitutionellt större.

Regionvariablerna är de som får de största utslagen i flera av modellerna. Precis som för *ethnic* har regionvariablerna mer effekt för de konstitutionella modellerna. Detta beror troligtvis på att region 4 som är jämförelsekategori (Europa, Nordamerika och Australien) har betydande övervikt från ledare som förlorar makten konstitutionellt. Afrika och Mellanöstern har betydligt längre överlevnadstider för ledare som förlorar makten konstitutionellt, även för materialet som helhet. Ledare som förlorar makten demokratiskt tycks alltså sitta mycket längre vid makten i Mellanöstern och Afrika, dessa regioner har även högre medelvärden för överlevnadstiden för icke konstitutionell maktförlust även om koefficienterna i den AFT modellen inte är signifikant. De icke signifikanta regionkoefficienterna för icke konstitutionell maktförlust kan bero på valet av region 4 som jämförelsekategori. Region 4 har endast 41 observationer av 741 (5,5 procent) som är icke konstitutionella.

Bienen och van de Walle (1991) tog med variabeln *multiple* för att utreda om ledare som redan varit vid makten en gång samlar kunskap och erfarenheter som gör att de lyckas hålla makten längre nästa tillfälle. Utifrån resultatet i tabell 4.9 verkar denna hypotes inte stämma. Ledare som är vid makten första gången har längre överlevnadstider och lägre risk för att förlora makten. Kanske kan ledare som redan varit vid makten lätt bli överexponerade, människor kan bli trötta snabbare på ledare som redan varit vid makten en gång. *Multiple* är dock inte signifikant för modellerna med icke konstitutionell maktförlust. En möjlighet är att variabeln *age* istället har fångat upp ledarens erfarenhet och kunskap, men detta visar sig inte vara fallet heller. Effekterna är mycket svaga för enskilda år men en äldre ledare tycks ha svårare att behålla makten än en yngre. Så även i detta fall är erfarenhet inte nödvändigtvis synonymt med längre tid vid makten. Resultatet är likadant för de båda händelserna men lite svagare för den icke konstitutionella modellen.

Literacy har precis som *age* svaga procentuella effekter på risk och överlevnadstid. Högre läskunnighet verkar innebära kortare tid vid makten och högre risk att förlora makten för konstitutionell maktförlust. Tvärtom gäller för icke konstitutionell maktförlust då högre läskunnighet resulterar i mindre risk och längre överlevnadstid för denna händelse. Resultatet stämmer överens med Bienen och van de Walles hypotes för konstitutionell maktförlust, men inte för icke konstitutionell maktförlust. Även modellering av region 4 (tabell D.4) visar på motsägande resultat då läskunnigheten verkar minska risken för konstitutionell maktförlust. *Literacy* är i datamaterialet en landsspecifik variabel som har ett konstant värde under hela mätperioden. Troligtvis skulle tilltron till resultatet bli något bättre om variabeln istället var ledarspecifik, läskunnigheten i landet skulle då vara uppmätt när ledaren regerar i landet istället för en ospecificerad tidpunkt som fallet är nu. Läskunnigheten förändras alltså inte för något land under hela mätperioden vilket är orealistiskt. De kvarvarande kontinuerliga variablerna *population* och *land* är visserligen signifikanta för flera av modellerna men verkar inte ha någon inverkan på risk eller överlevnadstid för såväl konstitutionell som icke konstitutionell maktförlust. Länder som är större till ytan och population tycks inte innebära någon ökad svårighet för en ledare att regera, vilket även säger emot Bienen och van de Walles hypoteser.

Ett av de stora problemen med datamaterialet är att vissa länder har ledare som regerar väldigt kort tid. Detta medför att länder med kortare överlevnadstider även tillför betydligt fler observationer då fler ledare hinner regera under studieperioden. Exempelvis har Mexico 72 observationer från tidigt 1800-tal till studieperiodens slut och Frankrike 81 observationer bara under 1900-talet medan Iran endast har 3 ledare från 1925 och framåt. Länder med korta överlevnadstider är alltså överrepresenterade. Detta gör att Latinamerika, som har flera länder med många ledare, är en majoritet av datamaterialet och tolkningarna som görs är till stor del baserat på just Latinamerika och till viss del även Region 4¹⁶. Många av variablerna är just landsspecifika vilket gör att länder med många ledare har större påverkan på just dessa variabler (exempelvis etnisk konflikt och läskunnighet). Samtidigt visar modellerna på att skillnaden är stor mellan regionerna vad gäller överlevnadstiden och risken för maktförlust. Denna kombination, stor skillnad mellan regionerna och att vissa regioner är överrepresenterade, gör att skattningarna blir svåra att generalisera för hela världens ledare.

Vad som gör att skillnaden är så stor mellan regionerna och att region som variabel är högt signifikant besvaras dock inte av denna studie. För det första är regionerna indelade med strävan efter homogenitet. Länder i Europa, Nordamerika och Australien kan antas ha mycket gemensamt likaså länder i Mellanöstern eller Latinamerika. Vad som egentligen är dessa ”gemensamma karaktärsdrag” som påverkar ledarnas överlevnad kan dock endast spekuleras om. Troligtvis kan det bero på socioekonomiska förutsättningar, graden av demokrati, kulturella aspekter och/eller ledarstilar för olika delar av världen med mera.

För att hantera dessa problem skulle för det *första* behövas metoder som tar hänsyn till att vissa länder är överrepresenterade i datamaterialet. Som det ser ut nu är skattningarna inte skattningar för hela världens ledare utan snarare ledare från Latinamerika och Region 4. För det *andra* är det kanske inte så intressant i sig själv att en region har längre överlevnadstider än en annan utan snarare vad som är anledningen till detta. Nya förklarande variabler skulle alltså behövas som fångar upp skillnaderna mellan regionerna.

¹⁶ Många länder i Mellanöstern, Afrika och Asien har uppnått självständighet ganska sent i mätperioden vilket ger dessa regioner ganska få observationer och datamaterialet som helhet mycket ojämnt fördelat över regioner.

En *tredje* reflektion är att variabler som exempelvis etnisk konflikt och kanske ledarens bakgrund inte nödvändigtvis direkt förklarar ledarens tid vid makten utan snarare vilket statskicks ett land har. Om detta var fallet så skulle tolkningen av variabeln etnisk konflikt bli mer förståelig – lägre etnisk konflikt resulterar i större chans för demokrati och det ger vidare större risk för konstitutionell maktförlust.

För det *fjärde* så är det inte bara länder som är överrepresenterade utan till viss del även vissa ledare. En ledare kan regera flera gånger och på så vis bidra med flera observationer. Frågan är hur detta påverkar resultatet och om dessa bör behandlas på annorlunda sätt. Samtidigt är det intressant att jämföra deras överlevnadstid med ledare som endast är en enda gång vid makten. Även här kan det finnas behov av metoder som kan hantera överrepresentation av vissa personer, men detta är inte ett problem av samma grad som länders överrepresentation.

5.2 Bristfälliga metoder för konkurrerande risker

En brist som är ett större problem än vissa länders/regioners överrepresentation är avsaknaden av metoder för att modellera konkurrerande risker. Som resonerats är det mycket tveksamt att anta att konstitutionell och icke konstitutionell maktförlust inte skulle vara beroende gentemot varandra. Detta i sin tur medför viss misstanke om bias för modeller där den ena händelsen hölls censurerad för att modellera den andra maktförlusten. Var det korrekt att hantera ledare som dog av naturliga orsaker som censur och är dessa verkligen inte informativa? Skattningarna har visats sig bli mer signifikanta för båda händelserna tillsammans då endast naturlig död hanterats som censur, vilket beror på att naturlig död troligtvis orsakas av helt andra kovariater. Dock är naturlig död en konkurrerande händelse och medför att ingen av de andra två händelserna kan inträffa samt att en del, troligtvis ingen avgörande del, information går förlorad.

Problemet är vidare att vi inte kan vara säkra för hur stor denna bias egentligen är för censurering av de olika händelserna, framförallt konstitutionell och icke konstitutionell maktförlust. Pepe och Moris test gav dock en viss fingervisning om att variabler som exempelvis ledarens bakgrund och hur ledaren tog makten är signifikanta då hänsyn tas till den konkurrerande risken, dessutom i samma riktning som skattningarna då hänsyn inte togs till detta i Wilcoxon testet. Det innebär att konkurrerande risk inte verkar orsaka något större problem för ledarens bakgrund och typ av makttillträde. Den etniska konflikten som dikotom variabel (låg etnisk konflikt kontra mellan och hög) var inte signifikant för icke-konstitutionell utgång vilket även varit fallet för Cox och AFT modellerna. Att resultatet är liknande för Pepe och Moris test som för Cox och AFT modellerna, som inte tar hänsyn till konkurrerande risker, gör att biasen inte nödvändigtvis behöver vara så avgörande att den helt kan ändra skattningarnas riktning. Vi kan dock inte säga något om hur de kontinuerliga variablerna påverkas av konkurrerande händelser.

Om det nu finns problem med konkurrerande risker varför då inte bara modellera maktförlusterna tillsammans som en enda händelse? Svaret på varför detta inte är lämpligt fick vi i vår logit-modell. Denna modell, trots misstänkt bias, visade på att det fanns klart signifikanta skillnader för hur variablernas koefficienter skiljer sig mellan konstitutionell och icke konstitutionell maktförlust. De variabler där detta inte var fallet var etnisk konflikt, landets storlek, antalet invånare, ledarens militära/civila bakgrund och om ledaren regerat mer än gång. För variablerna över regionerna och hur ledaren tog makten skilde sig skattningarna betydligt för de två typerna av maktförlust. Detta visar på behovet av att modellera för dessa händelser var för sig då det kan bli missvisande när händelserna modelleras som en

gemensam händelse. Problemet är dock avsaknaden av metoder för att göra detta. Test har kunnat göras för specifika händelser med Pepe och Mori, baserat på kumulativa incidenser och betingade sannolikheter, men inte hela modeller där alla variabler, både kategoriska och kontinuerliga, inkluderas som i Cox och AFT. Dessa händelsespecifika modeller har bias, troligtvis inte avgörande, men behovet finns för mer utvecklade modeller som tar hänsyn till detta dilemma.

5.3 Vilken modell?

Flera olika metoder har applicerats på datamaterialet för att kunna förklara ledares tid vid makten. En av de viktigaste uppgifterna med uppsatsen är att hitta en lämplig modell som passar datamaterialet och ger den mest tillförlitliga informationen. Tre huvudsakliga metoder (Cox regressionsmodell, AFT och en logit modell) samt ett antal olika tester och estimat har anpassats för att analysera ledares tid vid makten. Problem och brister har funnits för samtliga metoder. En nackdel med alla tre huvudsakliga metoder, som särskilt gäller denna uppsats, är att ingen klarar av att ta hänsyn till att olika händelser kan vara informativa gentemot varandra. När varje händelse modelleras var för sig har övriga händelser behandlats som censur vilket är givet att ingen händelse ger någon information om en annan händelse.

Cox regressionsmodell har ett grundantagande att kovariaterna ska vara proportionella över tid. För att kontrollera antagandet har tre metoder använts, LLS diagram, Schoenfeld residualer och skapandet av tidsberoende variabler. LLS diagramen har visat sig vara svåra att utläsa och ger mest en fingervisning om icke proportionalitet. Schoenfeld residualer ger p-värden för signifikansen och därigenom ett något mer definitivt svar, men är känsliga för valet av transformation för tidsaxeln samt antalet valda frihetsgrader. Transformationen av tiden är även ett problem för skapandet av tidsberoende variabler. All tre metoder kan dock tillsammans ge en bild om proportionalitetsantagandet för en variabel bör förkastas eller ej.

Att anpassa distributioner för överlevnadstiden i datamaterialet är inte okomplicerat eftersom de tillgängliga distributionerna inte riktigt kan ta hänsyn till den plötsliga ökningen av risk som sker efter fyra år vid makten. AFT modellernas skattningar blir då inte lika säkra som det skulle önskas. En fördel är att de olika distributionerna har kunnat testas som specialfall av en EGG modell (egentligen ett test av parameterfixeringar) men detta går inte att göra med Log-logistic distributionen eftersom den inte kan specificeras genom en EGG modell. Log-logistic distributionen får då istället testas grafiskt vilket har visat sig mycket godtyckligt och resulterat i att ingen Log-logistic modell har kunnat användas. Anpassningen av en Log-normal modell gav dock "bra" grafisk anpassning för icke konstitutionell maktförlust i figur C.6, dock visade sig skattningarna ha sämre signifikans i jämförelse med motsvarande Cox regressionsmodell för samma händelse. AFT-modeller sparar frihetsgrader då parametrar fixeras, vilket borde vara en fördel för Log-normal modellen. Detta visar på att det är godtyckligt och svårt att grafiskt bedöma om ett datamaterial har en specifik fördelning.

Multinomial logit-modellen är olik de två andra metoderna eftersom den modellerar skillnaden mellan betingade sannolikheter för de två händelserna konstitutionell och icke konstitutionell maktförlust utifrån skillnader i koefficienterna till de förklarande variablerna. Den är därför svårare att tolka och att jämföra med de två andra modellerna. Det ges heller inga skattningar för hur variablerna påverkar sannolikheten över tiden utan bara för en enskild tidsperiod. Detta kan också vara en fördel då det inte finns något problem med tidsberoende variabler. Skattningarna i logit-modellen gäller inte för de enskilda händelserna utan hur de förhåller sig till varandra, vilket gör att test kan utföras för signifikanta skillnader i hur variablerna påverkar de två händelserna konstitutionell och icke konstitutionell maktförlust.

Enligt ovan förda resonemang kan Cox regressionsmodell vara den mest lämpliga modell för datamaterialet. Detta beror på att den bäst kan hantera materialets utformning då ingen riktigt passande distribution gick att finna med AFT. I och med att Cox även gav möjligheten att hantera tidsvariation inom variablerna genom stratifiering och skapande av nya tidsberoende variabler, kunde en modell skapas utan att anta proportionell hazard. Vidare anses att händelserna konstitutionell och icke konstitutionell maktförlust bör hanteras var för sig eftersom viktiga förklarande variabler påverkar de två maktutgångarna mycket olika. Metoderna som tar hänsyn till att händelserna är informativa gentemot varandra har inte visat på några dramatiska skillnader när variablerna testats, men en viss bias antas finnas när den konkurrerande risken censureras för att analysera en händelse av intresse. Alltså trots att de valda Cox regressionsmodellerna för de två separata händelserna lider av bias på grund av beroendeförhållande anses modellernas fördelar väga tyngre än nackdelarna.

6 Slutord

Bienen och van de Walles (1991) stora slutsats i sin studie är att risken att förlora makten minskar med tiden. Distributionerna anpassade för hazarden har även visat på en stark nedgående trend för risken, samtidigt som en konstant hazard (exponentiell distribution) förkastats. Det finns dock en osäkerhet i slutsatsen om att risken minskar med tiden, vilket inte Bienen och van de Walle tagit hänsyn till. Den nedgående trenden i hazarden behöver inte nödvändigtvis bero på att risken minskar med tiden utan snarare att högrisk ledare förlorar makten tidigt i studien och sedan är det endast ledare med låg risk kvar i slutet av tidsperioden. Detta kallas för att det finns icke observerad heterogenitet i datamaterialet¹⁷. Specifika ledare kan ha konstanta nivåer på sin hazard, antingen hög eller låg, och när de ledare med hög risk försvinner ger det en hazardkurva som sluttar trots att risken egentligen inte minskat för ledarna som är kvar. För vissa länder, som exempelvis i Latinamerika, kan det möjligtvis vara så att hazarden är relativt konstant, men sedan ökar dramatiskt vid valår för att sedan gå tillbaka till samma nivå som tidigare på hazarden (vilket förklarar den tidigare spiken år 4 för datamaterialets hazardkurva). Denna uppsats kan inte bekräfta om detta resonemang gäller för materialet i fråga, men försiktighet bör beaktas vad gäller om risken *verkligen* minskar över tiden för ledare.

Utöver detta har denna uppsats visat på att det är svårt att anpassa distributioner för datamaterial med oregelbunden form på överlevnadstiden. Material av denna karaktär får bättre skattningar med Cox regressionsmodell, givet att tidsberoende inom variabler kan upptäckas och åtgärdas. Eftersom det finns stor skillnad mellan hur variabler påverkar konstitutionell och icke konstitutionell maktförlust var för sig så blir det missvisande att hantera dessa som gemensam händelse. Metoderna har dock brister vad gäller hantering av händelser som är informativa gentemot varandra, vilket i sin tur skapar bias i skattningarna. För detta material har däremot denna bias uppskattats som liten och informationsvinsten att hantera vardera händelse för sig har ansetts vara större. Slutligen är även vissa länder och regioner överrepresenterade i materialet vilket gör att skattningarna blir mindre lämpliga för regioner som bidrar med få observationer. Eftersom regionerna dessutom skiljer sig åt i många bemärkelser bör försiktighet tas då generaliseringar för hela världen används för att beskriva ledare i vissa regioner.

¹⁷ Vidare läsning om icke observerad heterogenitet finns att läsa i Vaupel och Yashin (1985) samt i Allison (1995).

En mer ingående analys av i vilken grad det finns icke observerad heterogenitet för materialet kan vara intressant för framtida studier. Vidare forskning bör även sträva efter att hitta modeller som bättre klarar av att hantera konkurrerande risker med beroende mellan händelserna. Metoder finns för att testa enskilda variabler och modellera kumulativa incidenser i viss utsträckning, men inga välutvecklade metoder som har kapacitet att ge skattningar för flera kategoriska och kontinuerliga variabler på samma gång. Datamaterialet är också väldigt ojämnt fördelat över regioner och länder. Eftersom regionvariabeln har visat starka samband med ledarnas överlevnad skulle uppsatsens resultat vara mer trovärdigt om observationerna var jämnare fördelat. Det kan därför vara av intresse att använda ett annat datamaterial eller bara delar av det befintliga för att förklara ledares överlevnad. En mer djupgående analys av regionerna var för sig kan alltså vara intressant. Likaså kan bättre variabler tas fram som förklarar det som regionvariabeln sammanfattar, som exempel kan läskunnighet och ekonomiska variabler vara mer anpassade för den tidsperiod ledaren är vid makten. Även variabler för religion, sysselsättning, urbaniseringsgrad, folkhälsa (medellivslängd, barndödlighet etc) och kanske valdeltagande är exempel på saker som kan påverka ledares tid vid makten som inte tas hänsyn till i Bienen och van de Walle (1991) och som dessutom inte behöver vara alltför problematiska att samla in.

Referenser

- Allison, P. (1995), *Survival Analysis Using the SAS[®] System: A Practical Guide*, Cary, North Carolina: SAS Institute Inc.
- Bergstrahl, E. Therneau, T.M. och Lennon, R. (2009), *Schoen SAS Macro*, Mayo Clinic, [<http://mayoresearch.mayo.edu/mayo/research/biostat/sasmacros.cfm>] –2009 05 20.
- Bienen, H. S. och van de Walle, N. (1991), *Of Time and Power: Leadership Duration in the Modern World*, Stanford, CA: Stanford University Press.
- Collett, D. (2003), *Modelling Survival Data in Medical Research*, Andra upplagan, London: Chapman & Hall.
- Cox, D.R. (1975), Partial Likelihood, *Biometrika*, 62, 269-276.
- Ghilagaber, G. (2005), The Extended Generalized Gamma Model and its Special Cases: Applications to Modeling Marriage Durations, *Quality and Quantity*, 39, 71-85
- Grambsch, P.M. och Therneau, T.M. (1994), Proportional hazards test and diagnostics based on weighted residuals, *Biometrika*, 81, 515-526
- Harrell, F.E. (1991), *Daspline SAS Macro*, Mayo Clinic, [<http://mayoresearch.mayo.edu/mayo/research/biostat/sasmacros.cfm>] – 2009 05 20.
- Islam, M.A. och Khondoker, M.R. (2009), Use of log-log survival function in modeling time-covariate interactions in Cox regression, *Journal of Statistical Planning and Inference*, 139, 1968-1973
- Kaplan, E.L. och Meier, P. (1958), Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, 457-481
- Lee, E. (1992), *Statistical Methods For Survival Data Analysis*, Andra Upplagan, New York: Wiley Interscience.
- Pintilie, M. (2006), *Competing Risks: A Practical Perspective*, Chichester: John Wiley & Sons.
- Pepe, M.S. och Mori, M. (1993), Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data?, *Statistics in Medicine*, 12, 737 – 751.
- Schoenfeld, D.A. (1982), Partial residuals for the proportional hazards regression model, *Biometrika*, 69, 239-241.
- Vaupel, J.W. och Yashin, A.I. (1985), Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics, *The American Statistician*, 39 no 3, 176-185.

Appendix A - Variabelbeskrivning

Years	Ledares tid vid makten beräknat i år (ental). Ledare som regerat mindre än ett år har kodats som 0.
Lost	Hur ledaren förlorade makten: 0 = fortfarande vid makten år 1987; 1 = förlorade makten på konstitutionell väg; 2 = förlorade makten genom att naturligt ha avlidit; 3 = förlorade makten på icke-konstitutionell väg.
Manner	Sättet ledaren tog makten på: 0 = tog makten på konstitutionell väg; 1 = tog makten på icke-konstitutionell väg.
Birth	Årtalet ledaren föddes.
Date	Året ledaren tog makten.
Age	Vilken ålder ledaren hade vid maktövertagandet (vilket i detta fall är variabeln <i>date</i> subtraherat med variabeln <i>birth</i>).
Military	Ledarens bakgrund: 0 = civil ledare; 1 = militär ledare.
Colonial	Kolonialbakgrund innan självständighet (endast kodad för Afrika, Mellanöstern och Asien): 0 = Storbritannien; 1 = Frankrike; 2 = Belgien; 3 = Portugal; 4 = Ingen; 5 = Nederländerna; 6 = Japan; 7 = Nya Zeeland; 8 = Australien; 9 = USA; 10 = Italien; 11 = Turkiet.
Ethnic	Den etniska konfliktnivån i landet: 0 = låg; 1 = medel; 2 = hög.
Order	I vilken ordning ledaren har makten för landet sedan självständighet eller mätperiodens början. För Afrika kodades dock en dummyvariabel: 0 = alla ledare efter den första ledaren; 1 = landets första ledare.
Income1	BNP per capita för ledarens land år 1973 mätt i samma årtals dollar.
Income2	BNP per capita för ledarens land år 1980 mätt i samma årtals dollar.
Growth	Genomsnittliga BNP-tillväxten per capita för ledarens land mellan 1965-1983.
Islam	Procentandel av landets invånare som är muslimer (endast kodat för Afrika och Mellanöstern); Procentandel av invånarna som är av europeisk härkomst (endast kodat för Latinamerika). (Årtal saknas.)
Region	Vilken region ledaren regerade i: 0 = Mellanöstern; 1 = Afrika; 2 = Asien; 3 = Latinamerika; 4 = Nordamerika, Europa och Australien.
Land	Landets areal i tusen kvadratkilometer.
Pop	Landets population i miljoner (årtal saknas).
Literacy	Nivån på läs- och skrivkunnigheten (årtal saknas).
Entry	Antalet gånger ledaren har suttit vid makten.
Multiple	Dummyvariabel för om ledaren suttit i makten mer än en gång: 0 = ledaren har endast varit i makten en gång; 1 = ledaren har varit vid makten minst två gånger.
Violent	Nivån på det politiska våldet direkt efter självständighet (endast kodat för Afrika, Asien och Mellanöstern); 0 = låg nivå; 1 = medium nivå; 2 = hög nivå.

Appendix B – Sammandrag av datamaterialet

Tabell B.1 *Frekvensdistribution för antalet år vid makten*

Antal år vid makten		Antal ledare (censur i parantes)	Antal ledare i procent	Kumulativa procenten
Intervall				
0	1	668 (13)	29,6	29,6
1	2	352 (18)	15,6	45,2
2	3	267 (15)	11,8	57
3	4	169 (12)	7,5	64,5
4	5	262 (7)	11,6	76,1
5	6	115 (15)	5,1	81,2
6	8	156 (19)	6,9	88,1
8	10	68 (17)	3	91,2
10	15	90 (15)	4	95,1
15	20	53 (14)	2,4	97,5
20	25	24 (11)	1,1	98,6
25	58	33 (9)	1,5	100
Totalt		2257 (165)	100	

Tabell B.2 *Antal observationer uppdelade efter händelse, typ av maktillträde och militär bakgrund*

LOST	MANNER				Totalt
	Konstitutionell		Icke Konstitutionell		
	MILITARY		MILITARY		
	Civil	Militär	Civil	Militär	
Kvar 1987	102	12	14	37	165
Konstitutionell	1030	165	78	121	1394
Naturlig död	84	24	11	13	132
Icke Konstitutionell	207	68	85	206	566
Totalt	1423	269	188	377	2257

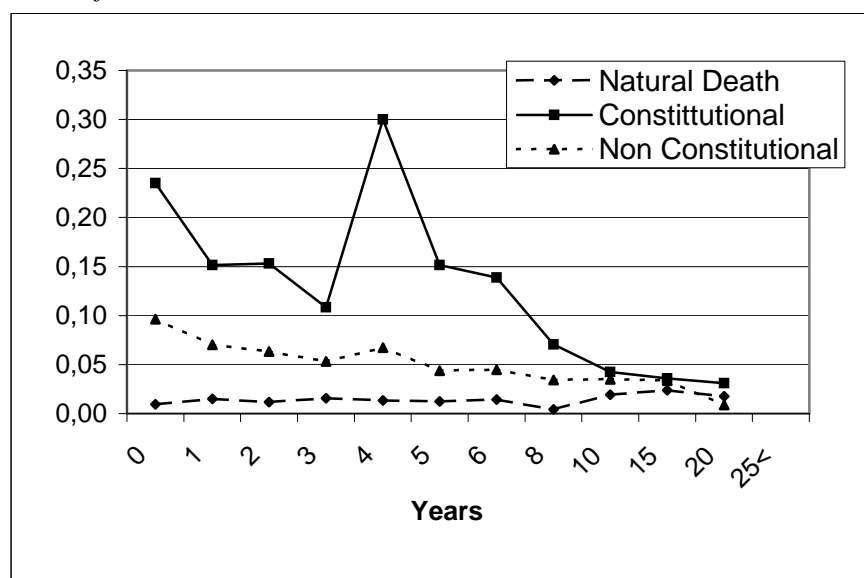
Tabell B.3 *Medelvärde och median för överlevnadstiden för de olika regionerna*

	Afrika	Mellanöstern	Asien	Latinamerika	Nordamerika, Europa och Australien	Hela datamaterialet
Medelvärde	7,03	7,43	4,85	2,96	2,7	3,57
Median	5	5	2	2	1	2

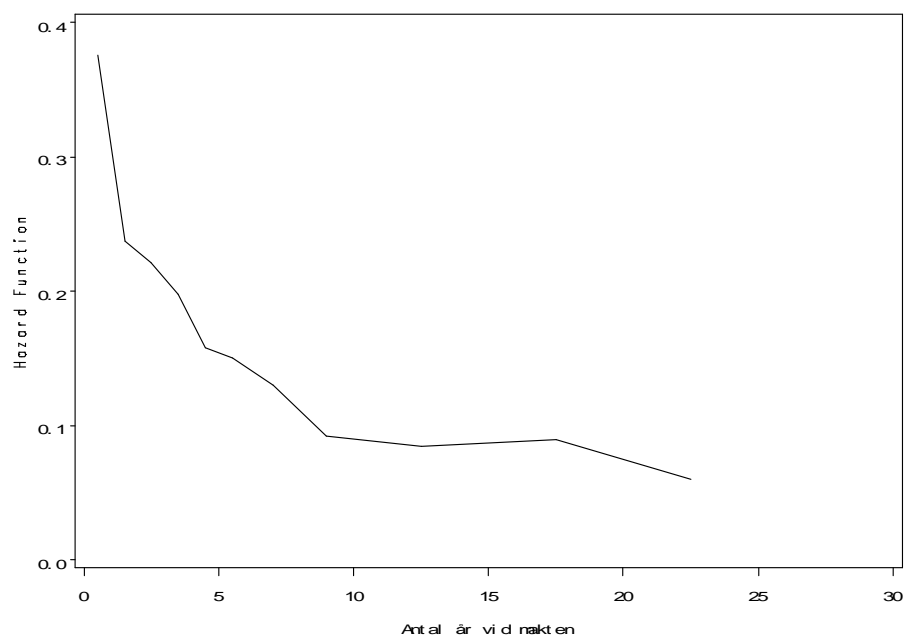
Tabell B.4 *Frekvensdistribution för antalet år vid makten för varje region (censur i parantes)*

Antal år vid makten		Regioner						Hela datamaterialet
Intervall		Afrika	Mellanöstern	Asien	Latinamerika	Nordamerika Europa och Australien		
0	1	24 (4)	18 (1)	52 (3)	266 (1)	308 (4)		668 (13)
1	2	14 (2)	12 (2)	40 (4)	164 (6)	122 (4)		352 (18)
2	3	16 (4)	13 (1)	30 (2)	129 (5)	79 (3)		267 (15)
3	4	14 (1)	5 (0)	25 (3)	71 (5)	54 (3)		169 (12)
4	5	6 (0)	9 (1)	8 (1)	199 (4)	40 (1)		262 (7)
5	6	8 (4)	6 (2)	15 (1)	55 (3)	31 (5)		115 (15)
6	8	11 (7)	14 (1)	10 (2)	77 (5)	44 (4)		156 (19)
8	10	17 (6)	5 (4)	12 (5)	17 (0)	17 (2)		68 (17)
10	15	20 (7)	11 (1)	14 (4)	23 (1)	22 (2)		90 (15)
15	20	9 (4)	13 (6)	12 (2)	8 (0)	11 (2)		53 (14)
20	25	9 (5)	1 (0)	7 (4)	3 (1)	4 (1)		24 (11)
25	58	2 (1)	8 (2)	6 (2)	7 (2)	9 (2)		33 (9)
Totalt		151 (45)	115 (21)	231 (33)	1019 (33)	741 (33)		2257 (165)

Figur B.1 *Hazarden för de olika händelserna med intervall som i tabell B.1*



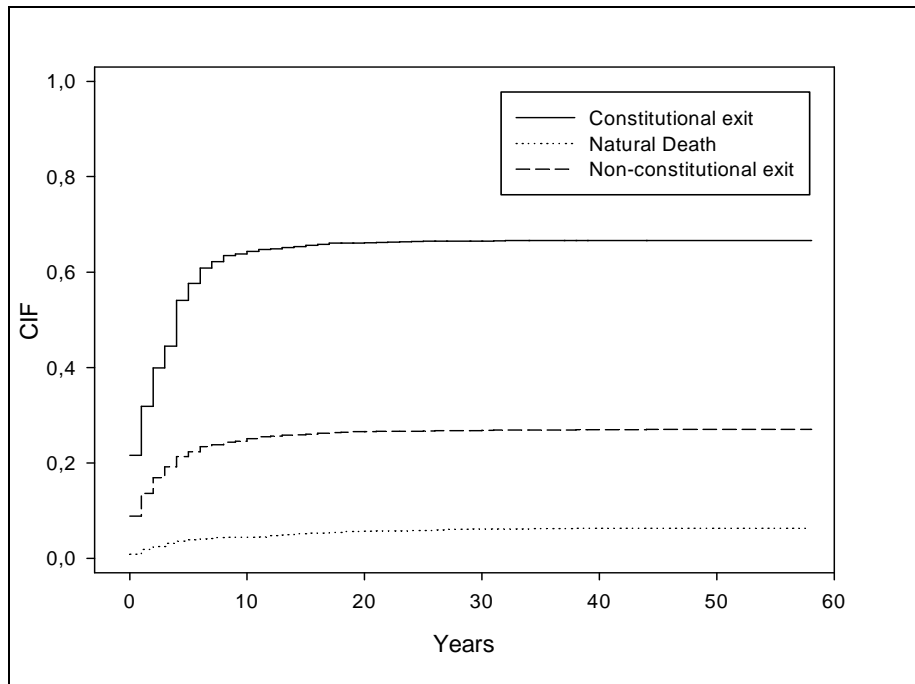
Figur B.2 Hazarden för hela materialet då Latinamerika har exkluderats



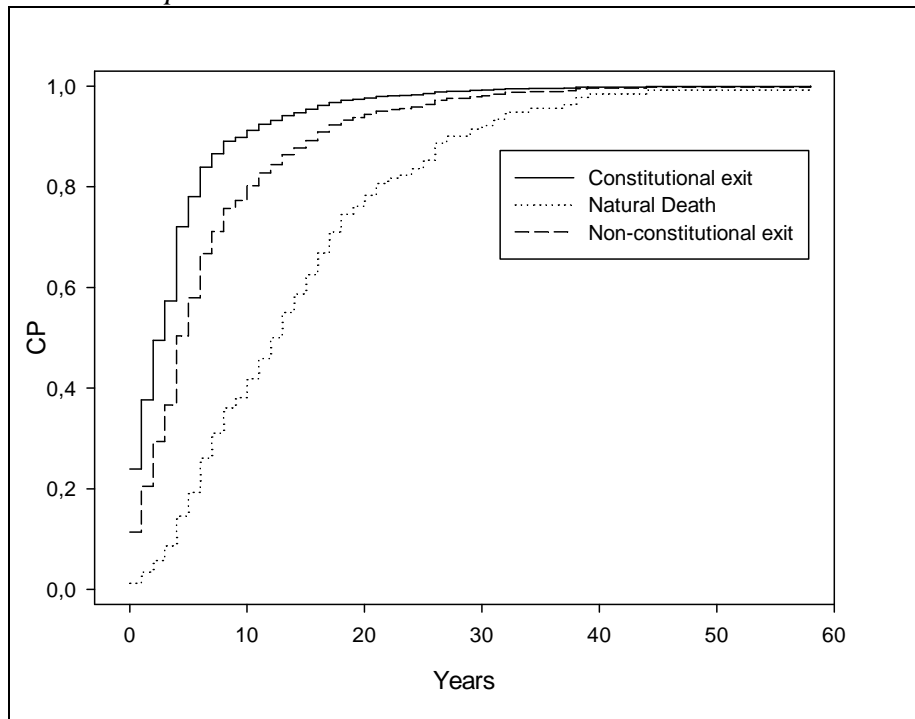
Tabell B.5 Antal observationer uppdelade efter händelse och region

LOST	REGION					Totalt
	Mellan- östern	Afrika	Asien	Latin- amerika	Nordamerika Europa Australien	
Kvar 1987	21	45	33	33	33	165
Konstitutionell	37	21	128	578	630	1394
Naturlig död	13	9	15	58	37	132
Icke Konstitutionell	44	76	55	350	41	566
Totalt	115	151	231	1019	741	2257

Figur B.3 Kumulativa incidensen för konstitutionell utgång, icke-konstitutionell utgång och naturlig död



Figur B.4 Sannolikheten att en specifik händelse inträffar givet att ingen annan händelse inträffar under samma tidpunkt



Appendix C Sammandrag av resultat

C.1 Icke parametriska metoder

Tabell C.1 Log-rank och Wilcoxon test över variablerna för konstitutionell maktförlust, med tillhörande chi-två och p-värde i parantes.

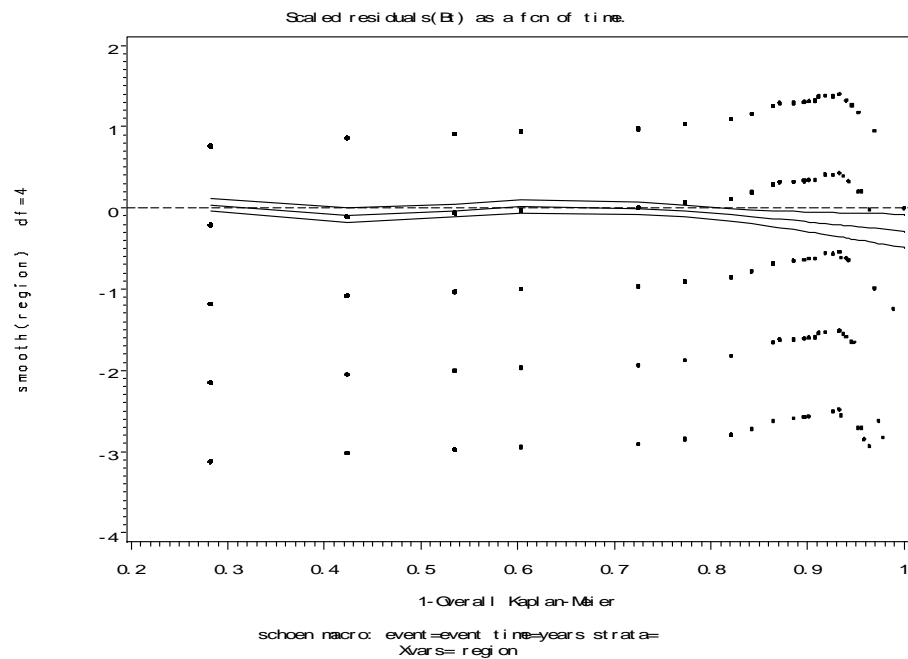
Variabel	Frihetsgrader	Log rank	Wilcoxon
Manner	1	86,61 (0,00)	49,39 (0,00)
Military	1	59,56 (0,00)	61,34 (0,00)
Ethnic	2	178,84 (0,00)	176,12 (0,00)
Region	4	387,30 (0,00)	373,80 (0,00)
Multiple	1	120,84 (0,00)	26,94 (0,00)

Tabell C.2 Log-rank och Wilcoxon test över variablerna för icke konstitutionell maktförlust, med tillhörande chi-två och p-värde i parantes.

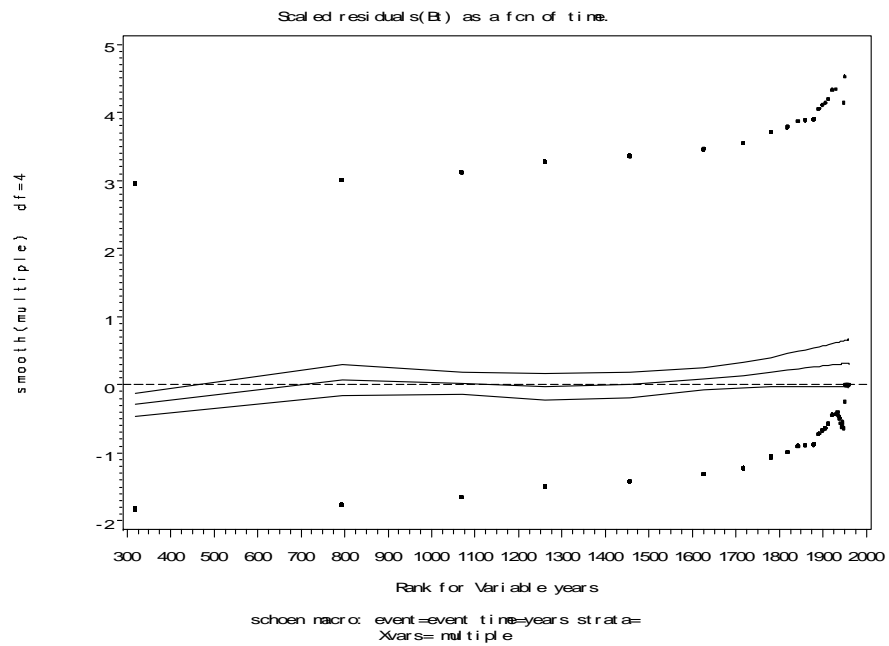
Variabel	Frihetsgrader	Log rank	Wilcoxon
Manner	1	247,27 (0,00)	266,33 (0,00)
Military	1	112,95 (0,00)	114,07 (0,00)
Ethnic	2	44,60 (0,00)	38,95 (0,00)
Region	4	155,77 (0,00)	120,27 (0,00)
Multiple	1	2,50 (0,11)	6,55 (0,01)

C.2 Proportionalitetstest

Figur C.1 Schoenfeld residualer för region

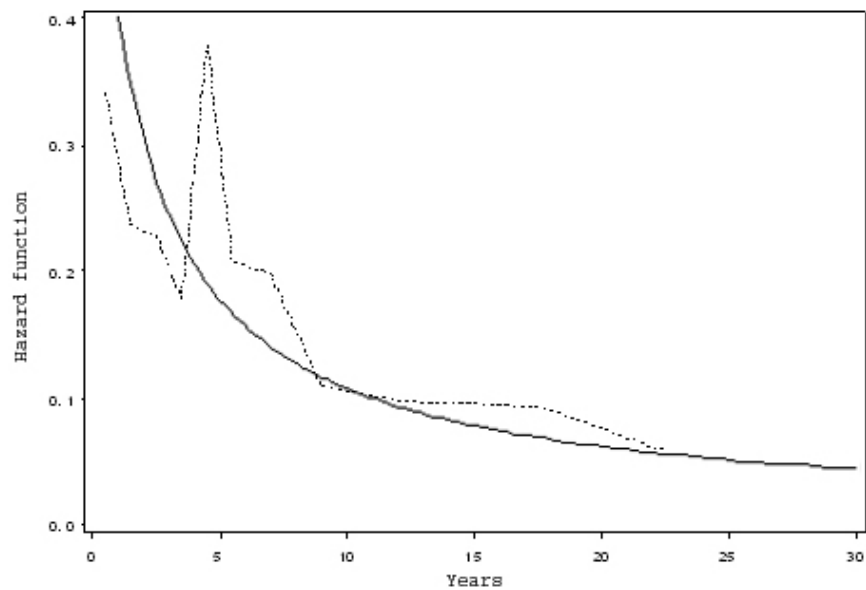


Figur C.2 Schoenfeld residualer för multiple

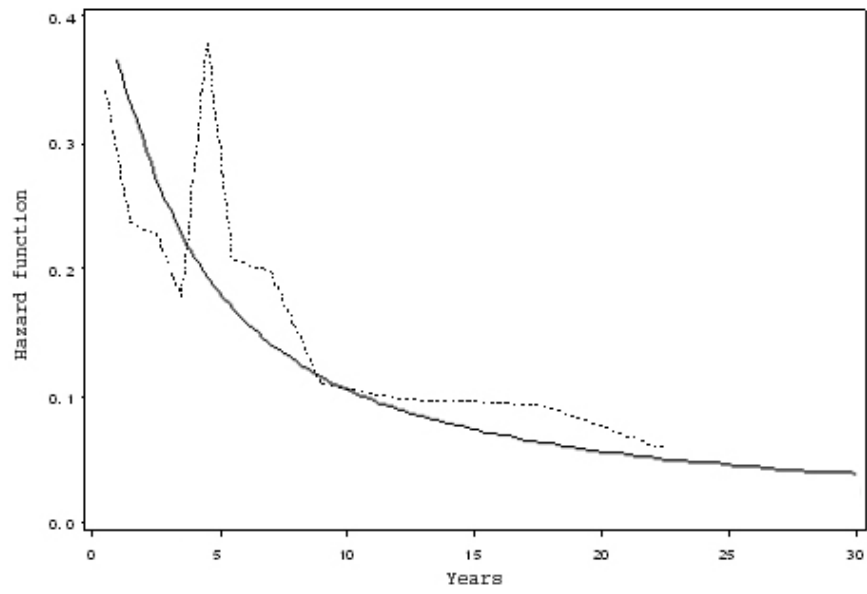


C.3 Distributionstest

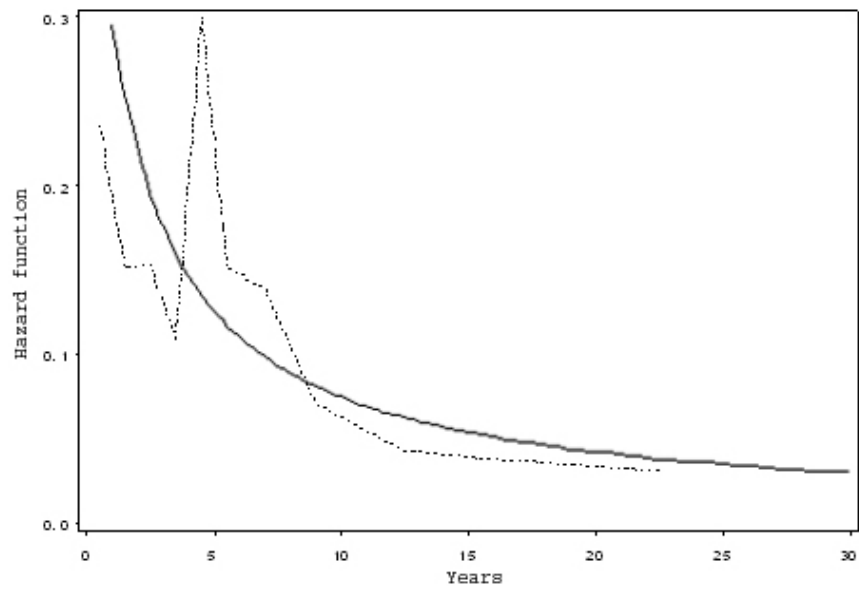
Figur C.3 EGG modellen anpassad till datamaterialets hazard



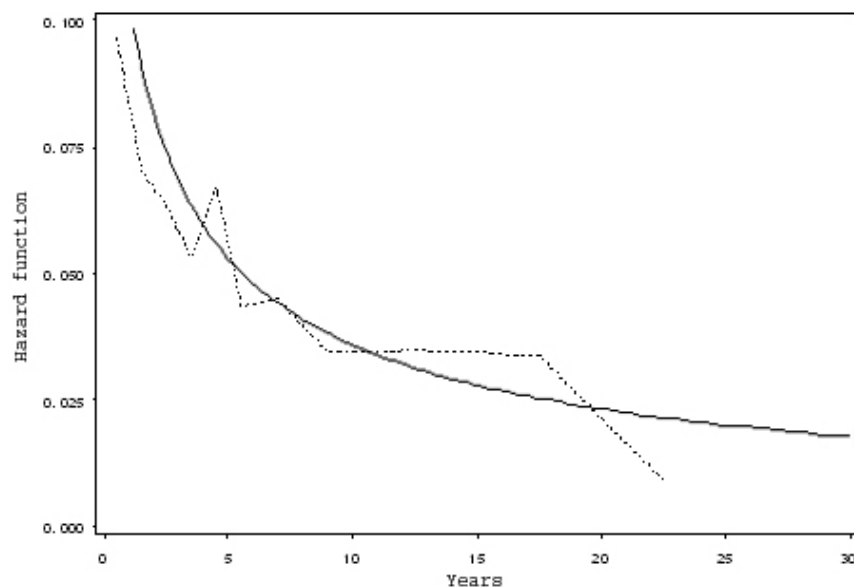
Figur C.4 *Log-logistic modell anpassad till datamaterialets hazard*



Figur C.5 *EGG modell anpassad till risken för konstitutionell maktförlust*



Figur C.6 Log-normal modell anpassad till risken för icke konstitutionell maktförlust



Tabell C.3 Log-likelihood test för för risken att förlora makten på konstitutionellt vis

Distribution	Log-likelihood	Chi-två
EGG	-2671,75	-
Log-Normal	-2677,19	10,88
Weibull	-2717,21	90,92
Gamma	-2814,27	285,04
Exponential	-2851,49	268,56
(Log-logistic)	(-2676,42)	-

Tabell C.4 Log-likelihood test för för risken att förlora makten på icke konstitutionellt vis

Distribution	Log-likelihood	Chi-två
EGG	-1581,93	-
Log-Normal	-1582,09	0,32
Weibull	-1587,87	11,88
Gamma	-1594,9	25,94
Exponential	-1645,74	115,74
(Log-logistic)	(-1585,45)	-

C.4 AFT modell anpassning

Tabell C.5 Resultat från AFT modell med konstitutionell maktförlust och EGG fördelning.

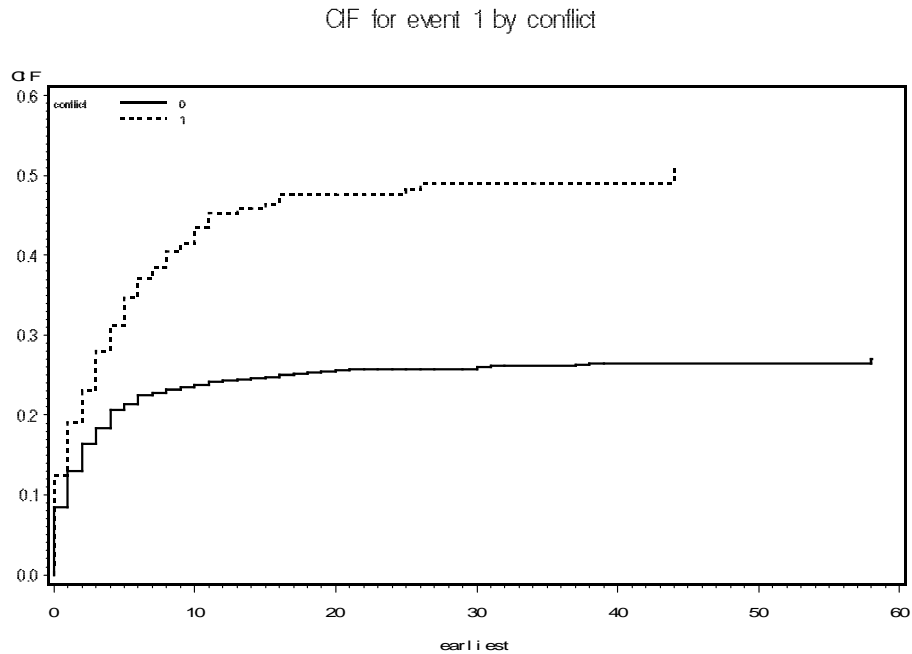
Variabler	Estimat	Standardfel	e^{β}	Chi-två	P-värde
Intercept	3,247	0,333		94,88	0,00
Manner (konstitutionell)	-0,098	0,106	0,907	0,86	0,35
Military (civil)	-0,079	0,094	0,925	0,70	0,40
Ethnic 0 (låg)	-0,356	0,137	0,701	6,78	0,01
Ethnic 1 (mellan)	-0,244	0,139	0,784	3,09	0,08
Region 0 (Mellanöstern)	1,413	0,206	4,109	46,94	0,00
Region 1 (Afrika)	1,985	0,251	7,279	62,36	0,00
Region 2 (Asien)	0,720	0,136	2,054	27,85	0,00
Region 3 (Latinamerika)	0,418	0,095	1,518	19,42	0,00
Multiple (första)	0,390	0,074	1,477	27,75	0,00
Age	-0,026	0,003	0,975	56,48	0,00
Literacy	-0,011	0,003	0,989	17,66	0,00
Population	-0,001	0,000	0,999	2,66	0,10
Land	0,000	0,000	1,000	18,01	0,00
Scale	1,291	0,033			
Shape	0,248	0,119			

Tabell C.6 Resultat från AFT modell med icke konstitutionell maktförlust och Log-normal fördelning.

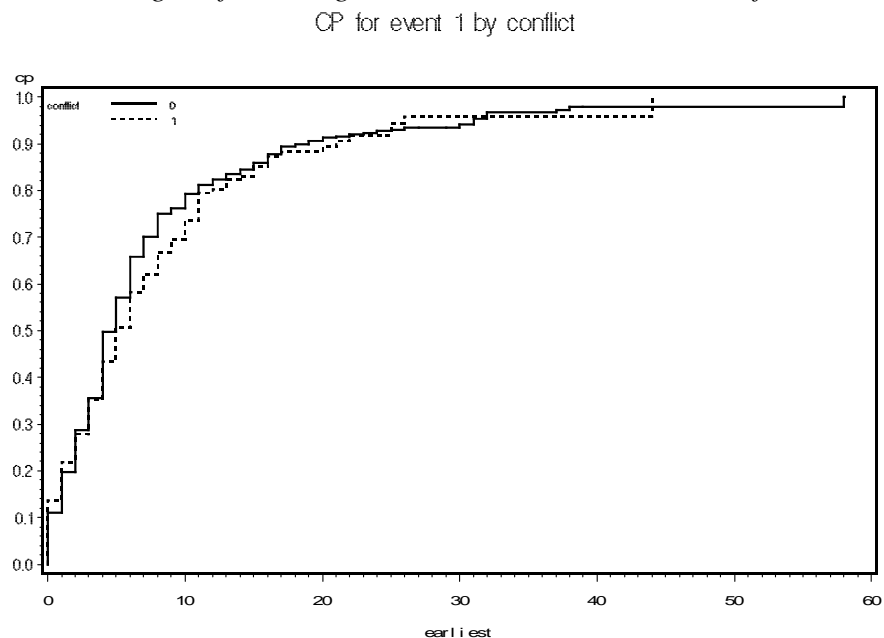
Variabler	Estimat	Standardfel	e^{β}	Chi-två	P-värde
Intercept	2,158	0,491		19,34	0,00
Manner (konstitutionell)	1,321	0,149	3,7464	78,98	0,00
Military (civil)	0,084	0,144	1,0880	0,34	0,56
Ethnic 0 (låg)	-0,017	0,200	0,9836	0,01	0,93
Ethnic 1 (mellan)	-0,246	0,187	0,7818	1,73	0,19
Region 0 (Mellanöstern)	-0,243	0,320	0,7846	0,58	0,45
Region 1 (Afrika)	-0,543	0,316	0,5813	2,95	0,09
Region 2 (Asien)	-0,007	0,291	0,9934	0,00	0,98
Region 3 (Latinamerika)	-1,213	0,199	0,2974	37,19	0,00
Multiple (första)	0,100	0,144	1,1048	0,48	0,49
Age	-0,015	0,006	0,9853	6,62	0,01
Literacy	0,011	0,004	1,0114	10,14	0,00
Population	-0,002	0,001	0,9984	5,13	0,02
Land	0,000	0,000	1,0001	4,04	0,04
Scale	1,838	0,076			
Shape	-	-	-	-	-

C.5 Kumulativ incidens

Figur C.7 Den kumulativa incidensen för icke-konstitutionell maktförlust för låg etnisk konflikt kontra mellan/hög konfliktnivå.



Figur C.8 Den betingade sannolikheten för icke-konstitutionell maktförlust för låg etnisk konflikt kontra mellan/hög konfliktnivå givet att inte konstitutionell maktförlust inträffat.



Appendix D Modeller över subgrupper

D.1 Separata modeller för regioner

Tabell D.1 Cox och AFT modeller för Afrika och Mellanöstern.

Variabler	Hela materialet		Konstitutionell		Icke konstitutionell	
	Hazard Ratio	AFT e^β (Weibull)	Hazard Ratio	AFT e^β (Gamma)	Hazard Ratio	AFT e^β (Log-normal)
Manner (konstitutionell)	0,57*	1,95*	0,80	1,33	0,45**	2,81**
Military (civil)	1,41	0,64	2,08	0,43	1,74	0,55
Ethnic 0 (låg)	0,84	1,36	0,43	2,24	0,85	1,03
Ethnic 1 (mellan)	0,95	1,11	0,36	2,93	1,22	0,71
Multiple (första)	0,41**	3,09**	0,36**	3,27**	0,55*	2,38*
Age	1,02**	0,98**	1,04**	0,96**	1,01	1,00
Literacy	1,00	1,00	1,03*	0,97*	0,99*	1,02*
Population	1,01*	0,99*	1,01	0,99	1,01*	0,99*
Land	1,00	1,00	1,00	1,00	1,00	1,00
Islam	1,00	1,00	1,02*	0,99*	0,99*	1,01
Koloni	0,64	1,5	0,79	1,45	0,43	2,35
Violent (låg)	1,134	0,88	1,03	1,01	1,65	0,84
Violent (mellan)	1,572	0,58	0,78	1,17	2,99**	0,35*

* = Signifikant på 5 procentnivå. ** = Signifikant på 1 procentnivå.

Tabell D.2 Cox och AFT modeller för Asien.

Variabler	Hela materialet		Konstitutionell		Icke konstitutionell	
	Hazard Ratio	AFT e^β (EGG)	Hazard Ratio	AFT e^β (EGG)	Hazard Ratio	AFT e^β (Log-normal)
Manner (konstitutionell)	1,24	1,18	2,38*	0,39*	0,74	1,58
Military (civil)	0,99	1,01	1,05	0,91	0,70	1,27
Ethnic 0 (låg)	1,14	1,01	1,18	0,84	0,43	2,83
Ethnic 1 (mellan)	1,05	0,85	0,73	1,44	2,59*	0,40
Multiple (första)	0,57**	1,57	0,59**	1,97**	0,54	1,84
Age	1,05**	0,94**	1,04**	0,95**	1,05**	0,94**
Literacy	1,00	1,00	1,01	0,99	0,98**	1,03**
Population	1,00	1,00	1,00	1,00	1,00	1,00
Land	1,00	1,00	1,00	1,00	1,00	1,00
Koloni	1,70*	0,79	1,90*	0,48*	0,79	1,40
Violent (låg)	1,21	0,65	1,00	0,91	1,69	0,39
Violent (mellan)	0,80	1,04	0,72	1,61	1,54	0,61

* = Signifikant på 5 procentnivå. ** = Signifikant på 1 procentnivå.

Tabell D.3 Cox och AFT modeller för Latinamerika.

Variabler	Hela materialet		Konstitutionell		Icke konstitutionell	
	Hazard Ratio	AFT e^β (EGG)	Hazard Ratio	AFT e^β (Log-normal)	Hazard Ratio	AFT e^β (Log-normal)
Manner (konstitutionell)	0,71**	1,65**	1,13	1,19	0,41**	3,36**
Military (civil)	1,08	0,94	1,21	0,87	0,92	1,05
Ethnic 0 (låg)	1,24	0,71*	1,44	0,59**	1,10	0,84
Ethnic 1 (mellan)	1,49**	0,64**	1,98**	0,51**	1,16	0,79
Multiple (första)	1,00	0,92	0,89	0,99	1,18	0,81
Age	1,01*	0,99*	1,01*	0,99*	1,00	0,99

Literacy	1,00	1,00	1,01*	0,99	1,00	1,01
Population	1,01	0,99**	1,00	0,99*	1,01	0,98*
Land	1,00	1,00**	1,00	1,00	1,00	1,00*
European	1,00	1,00	1,00	1,00	1,00	1,00

* = Signifikant på 5 procentnivå. ** = Signifikant på 1 procentnivå.

Tabell D.4 Cox och AFT modeller för Nordamerika, Europa och Australien.

Variabler	Hela materialet		Konstitutionell		Icke konstitutionell	
	Hazard Ratio	AFT e^β (EGG)	Hazard Ratio	AFT e^β (EGG)	Hazard Ratio	AFT e^β (Gamma)
Manner (konstitutionell)	0,76	1,43	1,18	1,01	0,17**	59,02**
Military (civil)	1,35	0,65	1,90**	0,46*	0,41	1,66
Ethnic 0 (låg)	2,36**	0,43*	1,94*	0,44*	-	-
Ethnic 1 (mellan)	1,32	0,90	1,22	0,91	-	-
Multiple (första)	0,64**	1,55**	0,63**	1,61**	0,67	2,61
Age	1,02**	0,98**	1,02**	0,98**	1,04*	0,89**
Literacy	0,96**	1,08**	0,97**	1,08**	0,91**	1,28**
Population	1,00*	1,00	1,00	1,00	1,01	0,98
Land	1,00*	1,00*	1,00**	1,00*	1,00	1,00

I dessa modeller har några nya variabler inkluderats som är specifika för vissa regioner. För det första finns en variabel *islam* för procentuella antalet muslimer i Afrika och Mellanöstern samt en variabel *european* för procentuella antalet invånare med europeisk härkomst i Latinamerika. Dessa variabler visade dock på bristande signifikans och låg grad av påverkan. En dummyvariabel introduceras för om ett land har kolonial bakgrund eller inte. Denna variabel är endast signifikant för konstitutionell maktförlust i Asien där risken var 90 procent högre för ledare som regera över länder utan kolonial bakgrund. Slutligen introducerades en variabel *violent* för graden av politisk konflikt efter självständighet i Afrika, Mellanöstern och Asien. Variabeln visade sig endast vara signifikant för icke konstitutionell maktförlust i Afrika och Mellanöstern där det var nästan 200 procent högre risk för ledare i länder där konflikten var mellan i jämförelse med hög.

Det finns inga skattningar för den etniska konflikten för icke konstitutionell maktförlust i region 4. Detta beror på att antalet ledare med mellan och hög nivå var väldigt få. Annars var *ethnic* främst signifikant för konstitutionell maktförlust i Latinamerika och region 4. I Latinamerika var risken betydligt högre för ledare i länder med mellan nivå av konflikt att förlora makten, medan risken var högre för låg konfliktnivå i region 4 i jämförelse med hög nivå. Variabeln *military* blev signifikant för konstitutionell maktförlust i region 4, där risken är 90 procent högre att förlora makten för civila ledare jämfört med militära. Läskunnigheten hade även motsatt påverkan i region 4 jämfört med övriga regioner för konstitutionell maktförlust då högre läskunnighet resulterade i mindre risk. De övriga variablerna har liknande påverkan genom regionerna med mer eller mindre styrka.

D.2 Separat modell för 1960 - 1987

Tabell D.5 Cox och AFT modeller för observationer efter 1960.

Variabler	Hela materialet		Konstitutionell		Icke konstitutionell	
	Hazard Ratio	AFT e^b (Log-Normal)	Hazard Ratio	AFT e^b (Log-Normal)	Hazard Ratio	AFT e^b (Log-Normal)
Manner (konstitutionell)	0,78	1,63**	1,32	0,98	0,46**	2,91**
Military (civil)	1,07	0,88	0,88	0,99	1,62*	0,66
Multiple (första)	0,63**	1,57**	0,64**	1,64**	0,71	1,65
Ethnic 0 (låg)	1,10	0,98	1,34	0,77	0,70	1,53
Ethnic 1 (mellan)	1,13	0,91	1,06	1,00	1,17	0,79
Age	1,02**	0,98**	1,03**	0,97**	1,01	0,99
Literacy	1,00	1,00	1,01	0,99**	1,00	1,00
Population	1,00	1,00	1,00	1,00	1,00	1,00
Land	1,00	1,00	1,00	1,00	1,00	1,00
Mellanöstern	0,44**	2,50**	0,42**	2,50**	3,31*	0,42
Afrika	0,33**	3,78**	0,18**	6,86**	2,04	0,82
Asien	0,38**	3,27**	0,52*	2,47**	1,22	1,52
Latinamerika	0,72*	1,64*	0,73	1,56*	4,38**	0,31*
Growth	0,98	1,02	1,01	0,98	0,96	1,08
Log-Income1	0,84*	1,32**	0,99	1,18	0,60**	1,88**

* = Signifikant på 5 procentnivå. ** = Signifikant på 1 procentnivå.

Förutom att datamaterialet analyserades för enskilda regioner har även en separat analys gjorts för alla observationer från och med 1960. Anledningen till detta är att observationerna är jämnare fördelade över regionerna för denna tidsperiod, vilket gör att Latinamerika inte blir lika dominant som tidigare. Det finns också några ekonomiska variabler som kan användas för denna tidsperiod som kan visa sig signifikanta, dessa är *growth* och *Log-income1*. *Growth* är den genomsnittliga BNP tillväxten per capita mellan 1965 – 1983. *Log-income1* är logaritmerad BNP per capita för 1973. I listan över variabler i Appendix A finns även 1980 års BNP per capita tillgänglig som variabel, men 1973 bedöms som bättre eftersom den ligger mer i mitten av mätperioden 1960-1987.

Många av resultaten i tabell D.5 stämmer överens med de tidigare resultaten som sammanfattades i tabell 4.9, där hela tidsperioden inkluderades. *Military* är visserligen signifikant för icke konstitutionell maktförlust men då bara på 5 procentnivån för Coxmodellen. Det skulle innebära att civila ledare har cirka 60 procent högre risk att förlora makten icke konstitutionellt i jämförelse med militära ledare. *Latinamerika* är något mindre signifikant än tidigare för flera av modellerna. Det finns alltså inte lika tydliga skillnader i risk och överlevnadstid mot region 4 i tidsperioden efter 1960 i jämförelse med de tidigare modellerna med alla tidsobservationer. Ett undantag är att hazard rationen för icke konstitutionell maktförlust nu är 4,38 för *Latinamerika* jämfört med det tidigare 3,62. Av de nya variablerna verkar det bara vara *Log-income1* som tillför något till modellerna och då enbart för de två modellerna med icke konstitutionell maktförlust. Högre BNP verkar sänka risken (ge längre överlevnadstider) för händelsen icke konstitutionell maktförlust.