

Mathilde Elisabeth Mortvedt

Befolkning og intern konflikt

Hvorfor opplever befolkningsrike land oftere utbrudd av intern konflikt enn andre?

Masteroppgave i Lektor i samfunnsfag

Veileder: Halvard Buhaug

Juni 2020

Mathilde Elisabeth Mortvedt

Befolkning og intern konflikt

Hvorfor opplever befolkningsrike land oftere utbrudd av intern konflikt enn andre?

Masteroppgave i Lektor i samfunnsfag
Veileder: Halvard Buhaug
Juni 2020

Norges teknisk-naturvitenskapelige universitet
Fakultet for samfunns- og utdanningsvitenskap
Institutt for sosiologi og statsvitenskap



Kunnskap for en bedre verden

Sammendrag

Denne studiens hovedformål har vært å skape en dypere forståelse av hvorfor befolkningsrike land oftere opplever utbrudd av intern konflikt enn andre. Temaet er lite forsket på, selv om sammenhengen mellom konflikt og befolkningsstørrelse viser seg å være ett av få robuste resultater i litteraturen og forskningen. Denne analysen baserer seg på tanken om at det er andre faktorer som har en høy korrelasjon med befolkning som kan være med på å forklare dette forholdet. Fire ulike aspekter har blitt testet: (1) geografisk størrelse og avstand, (2) etnisk mangfold, (3) en større masse å rekruttere fra, og (4) økonomiske insentiver og korrupsjon. Analysen baserer seg på 6 tilhørende hypoteser. Resultatet av de kjørte regresjonsanalysene finner belegg for å anta at både geografisk størrelse, antall etniske relevante grupper og antall etniske ekskluderte grupper er med og påvirker forholdet mellom befolkningsstørrelse og konflikt. Videre viser regresjonsanalysene at det finnes belegg for at både størrelse på rural befolkning og størrelse på BNP påvirker den aktuelle sammenhengen, men her er resultatene noe mer usikre. Antall unge menn viser seg å være vanskelig å bruke som indikator, ettersom denne korrelerer sterkt med befolkningsstørrelse, og gir derfor uklare svar som er vanskelige å tolke.

Summary

This paper is written because of the puzzle that connects onsets of civil conflict and population size. The theme for this paper has not been investigated as much as other themes that concerns civil conflict, despite the robust empirical findings in the literature that connects conflict to bigger populations. This analysis is based on the thought that other factors that describe a country and correlate with population can influence the relationship between population and conflict. Four different aspects have been tested: (1) geographical size and distance, (2) ethnic diversity, (3) a bigger mass to recruit from, and (4) financial incentives and corruption. The analysis is based on 6 associated hypotheses. The result of the regression analysis shows that both geographical size, number of ethnic relevant groups and number of ethnic excluded groups has a significant impact on the relationship between civil conflict and population size. The regression analysis also shows that the size of the rural population and the total GDP of the country has an impact on the relationship, but these results are much more uncertain compared to the other factors. It also shows that the variable that measure young men is not an appropriate measure to use with population size, because of the too high correlation between these variables. This makes it hard to say whether the size of young men in a country also affects the relationship between civil conflict and population size.

Forord

Denne masteroppgaven symboliserer slutten på fem fine år i Trondheim og på lektorstudiet. Faglig har årene både vært spennende og utfordrende! Gjennom arbeidet med masteroppgaven har jeg sett viktigheten av å være både kildekritisk og objektiv som forsker. Som samfunnsfagslærer er dette viktige egenskaper som jeg vil jobbe for å utvikle hos mine fremtidige elever. Oppgavens tema har også gitt meg god kompetanse til å undervise om både gamle og nyere konflikter. Tusen takk til ISS og ILU for alt jeg har fått muligheten til å lære og ta del i, i løpet av disse fem årene.

Den største takken går imidlertid til veilederen min, Halvard Buhaug. Høsten 2019 inspirerte du meg til å skrive masteroppgave om konflikt gjennom faget «Årsaker til krig». Som veileder har du vært stødig og alltid gitt meg nyttige tilbakemeldinger. Du har hjulpet meg med både små og store ting, og oppgaven hadde ikke sett slik ut som den gjør i dag, hvis det ikke hadde vært for deg.

Til slutt vil jeg takke alle medstudenter i MLSAM2015 for artige påfunn, faglige (og ikke-faglige) diskusjoner, og støtte på alle plan. Uten dere hadde Trondheim virkelig ikke vært det samme! Jeg er så glad for at vi har vært en klasse, og for alle gode vennskap som har blitt til mens årene har gått. Tusen takk! Familie og andre venner fortjener også å bli nevnt, da dere alltid har hatt troen på meg, og bidratt med både støtte og motivasjon ikke bare i disse fem årene, men i livet generelt. Mina, jeg er også utrolig takknemlig for ditt engasjement og entusiasme! Nå venter (endelig) voksenlivet på den andre siden!

Trondheim, juni 2020
Mathilde Mortvedt

Innholdsfortegnelse

1. Innledning.....	1
2. Teorigjennomgang.....	5
2.1 Geografisk størrelse og avstand	6
2.2 En heterogen befolkningsgruppe	9
2.3 En større masse av potensielle opprørere	14
2.4 Store gevinster å hente ved å styre.....	18
2.5 Alternative teoretiske grunnlag	21
3. Data og metode.....	23
3.1 Gjennomgang av data	23
3.2 Avhengig variabel(Y) – Utbrudd av konflikt.....	23
3.3 Uavhengige variabler (X)	25
3.4 Kontrollvariabler.....	29
3.5 Deskriptiv statistikk	30
3.6 Metode.....	30
4. Analyse	41
4.1 Modeller kjørt på hele datasettet (1961-2017)	41
4.2 Kurvelineært forhold mellom konflikt og etniske grupper	45
4.3 Modeller kjørt på datasettet uten Kina og India	49
4.4 Samlet forklaringskraft	52
4.5 Sammenfattende analyse.....	53
5. Konklusjon	60
6. Litteraturliste.....	62
7. Appendiks.....	71
7.1 Liste over land-år kodet som intern konflikt i Onset-datasettet, men ikke i Armed Conflict-datasettet.....	71
7.2 Test for multikollinearitet for modellene 1-8 (VIF)	71
7.3 Betydningsfulle enheter med cooks D for modell 1.....	73
7.4 Liste over uteliggere med standardiserte residualer større enn 3 for modell 1.....	75
7.5 Modeller uten de to mest betydningsfulle enhetene	81
7.6 Modeller kjørt på datasettet delt i to tidsserier (1961-1990 og 1991-2017).....	82

Figurer

Figur 1: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert totalt areal i 2017.	7
Figur 2: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og antall relevante etniske grupper i 2017.	10
Figur 3: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og antall ekskluderte etniske grupper i 2017.	10
Figur 4: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert unge menn i 2017.	15
Figur 5: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert rural befolkning i 2017.	15
Figur 6: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert total BNP i 2017.	19
Figur 7: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og korrupsjon i 2017.	19
Figur 8: lvr2plot over uteliggere og betydningsfulle enheter.	39
Figur 9: Marginsplot av sammenhengen mellom konflikt og befolkning i modell 1.	42
Figur 10: Marginsplot som viser sammenhengen mellom konflikt og antall relevante etniske grupper.	48
Figur 11: Marginsplot som viser sammenhengen mellom konflikt og antall ekskluderte grupper.	49

Tabeller

Tabell 1: Oversikt over konfliktrelaterte dødsfall i de mest befolkede landene i 2018 (Data fra FN og UCDP).....	5
Tabell 2: Oversikt over konfliktrelaterte dødsfall i de minst befolkede landene i 2018 (Data fra FN og UCDP).....	6
Tabell 3: Deskriptiv statistikk for variablene	30
Tabell 4: Korrelasjonsanalyse	38
Tabell 5: Modell 1-8, 1961-2017.....	45
Tabell 6: Modell 3+4, kurvelineært forhold.....	47
Tabell 7: Modell 9-16, uten Kina og India.....	51
Tabell 8: Modell 17, fullmodell.....	53

1. Innledning

I litteraturen om konflikt finnes det flere empiriske resultater som ikke er robuste. Det er store variasjoner i resultater i ulike studier, blant annet fordi det blir brukt ulike definisjoner av borgerkrig, ulike tidsperioder og ulike datasett (Hegre & Sambanis, 2006). Et par studier tar opp denne problemstillingen, blant annet Hegre og Sambanis (2006) og Sambanis (2004), som forsøker å teste hvorvidt de mest diskuterte og brukte variablene gir de samme resultatene uavhengig av hvilket datamateriale som er benyttet. Hovedfunnene i disse analysene viser at det er få variabler som holder seg signifikante og som gir robuste resultater uavhengig av operasjonaliseringer og datasett. Variablene som skiller seg ut er mål på befolkningsstørrelse, økonomisk tilgang og politisk ustabilitet (Sambanis, 2004:847; Hegre & Sambanis, 2006:509). Uavhengig av operasjonalisering og datagrunnlag, viser disse tre variablene seg å gi omtrentlige samme resultater. Man kan derfor si at det er konsensus på feltet om at risikoen for borgerkrig er større i befolkningsrike, fattige og politisk ustabile land sammenliknet med andre.

Ett av de mest robuste empiriske funnene i litteraturen om intern konflikt viser altså til at befolkningsrike land oftere opplever borgerkrig enn mindre befolkede land (Fearon & Laitin, 2003; Collier & Hoeffler, 2004; Hegre & Sambanis, 2006). Det er imidlertid lite forsket på, og det finnes liten enighet om hvorfor man ser denne sammenhengen (Sambanis, 2004:26). Er det slik at befolkningsrike land oftere opplever konflikt fordi det er større uenighet blant innbyggerne? Er menneskene i folkerike land mer misfornøyd sammenliknet med de i mindre befolkede land? Er det flere å dele ressursene med? Er det vanskeligere å styre så mange mennesker? Er det lettere å rekruttere opprørere når det er så mange å ta av? Er det mer å tjene på å styre en befolkningsrik stat? Eller er det rett og slett slik at hvert enkelt menneske har en konstant sannsynlighet for å gjøre opprør, og at store land opplever mer konflikt fordi de består av et større totalt antall mennesker? Dette er alle relevante spørsmål å stille for å forstå sammenhengen mellom befolkning og konflikt. Ikke minst er dette viktig for å få en dypere forståelse for konflikt generelt.

Raleigh og Hegre sin analyse fra 2009, er en av svært få artikler som anerkjenner at sammenhengen mellom befolkning og konflikt er et «puzzle» som trenger å bli undersøkt. I analysen gjennomførte de en undersøkelse om hvordan tre grupper av størrelsesrelaterte forklaringer fungerer på intern konflikt. Disse tre gruppene tar opp tre ulike aspekter ved

gruppestørrelsen: (1) hver borger har en konstant sannsynlighet for å bli involvert i opprør, (2) heterogenitet og militær kapasitet grunnet distanse øker sannsynligheten for å bli involvert i opprør, og (3) befolkningstetthet og konsentrasjon øker sannsynligheten for å bli involvert i opprør (Raleigh & Hegre, 2009:3). Analysen, som er gjennomført på disaggregerte data for 14 afrikanske land, finner empiriske bevis som støtter alle de tre aspektene.

Raleigh og Hegre (2009) er den eneste artikkelen det er mulig å finne med et litteratursøk der forskningen omhandler sammenhengen mellom befolkning og konflikt helt konkret. Det er gjort et minimalt arbeid i å finne ut av hvorfor befolkningsrike land opplever mer konflikt enn andre, og det kan virke som at spørsmålet har blitt forlatt uten et skikkelig svar. Artikkelen til Sambanis (2004) og Hegre & Sambanis (2006) er begge rundt 15 år gamle, som vil si at det er nokså lenge siden påvirkningen befolkningsstørrelse har på konflikt ble kjent som et robust resultat. Raleigh og Hegre (2009) kan også plasseres i samme tidsperiode. Til tross for at befolkning er en mye brukt kontrollvariabel i forskningen på konflikt, og et fravær av denne variabelen i enhver analyse vil stille spørsmål til forskeren, er det altså et tema som er forsket lite på i seg selv. Det kan virke som at miljøet tar denne sammenhengen for gitt, har gått videre, og latt spørsmålet stå ubesvart. Selv om det er vanskelig å finne artikler som tar opp sammenhengen mellom befolkning og konflikt direkte, er det likevel noen artikler fra nyere tid som er innom temaet i større eller mindre grad. Acemoglu, Fergusson og Johnson (2019) undersøker for eksempel hvordan befolkningsvekst påvirker sannsynlighet for konflikt, og begrunner sammenhengen med knapphet. Noe forskning som omhandler det samme temaet blir altså fortsatt gjort, men da særlig på ulike typer befolkningsvekst, og ikke befolkningsstørrelse i seg selv.

Med utgangspunkt i Raleigh og Hegre (2009), ønsker jeg i denne oppgaven å bygge videre på deres funn, og undersøke «puzzelet» nærmere på landnivå med et større antall land inkludert. Oppgavens forskningsspørsmål er derfor «*Hvorfor opplever befolkningsrike land oftere utbrudd av intern konflikt enn andre?*». To av aspektene Raleigh og Hegre inkluderte i sin analyse er videreført i denne oppgaven, samtidig som det er lagt til andre alternative aspekter som kan forklare hvorfor befolkningsrike land oftere opplever intern konflikt enn mindre land. Aspektet som er utelatt fra analysen, konsentrasjon og befolkningstetthet, har blitt utelatt på grunn av manglende korrelasjon mellom variablene som operasjonaliserer dette aspektet og befolkningsstørrelse. Dette vil bli kommentert ytterligere i neste del.

Før oppgavens struktur vil bli presentert, er det viktig med en konkretisering av begrepet «konflikt». Konfliktbegrepet som brukes i denne oppgaven, er hentet fra «UCDP/PRIO armed conflict dataset». Intern konflikt forstås som «en uoverensstemmelse over styresett og/eller territorium, der det blir brukt væpnet makt mellom to parter der en av sidene er staten, og som resulterer i minst 25 kamprelaterte dødsfall innenfor ett kalenderår» (Gleditsch, Wallensteen, Eriksson, Sollenberg & Strand, 2002; egen oversettelse). Siden oppgaven omhandler intern konflikt, er hendelsene som inkluderes kjennetegnet av at de enten er konflikter mellom en stat og en eller flere interne opprørsgrupper, eller konflikter mellom en stat og en eller flere interne opprørsgrupper med innblanding fra andre stater på enten én eller begge sider. All bruk av opprør, borgerkrig eller intern konflikt videre i oppgaven viser til denne definisjonen.

Det følgende kapittelet tar for seg de teoretiske argumentene som danner grunnlag for de ulike hypotesene som senere skal testes. Egenskapene som presenteres som mulige forklaringer på sammenhengen mellom befolkningsstørrelse og konflikt er areal, antall relevante etniske grupper, antall ekskluderte etniske grupper, antall unge menn, størrelse på rural befolkning, grad av korrupsjon og størrelse på BNP. Det presenteres seks ulike hypoteser som alle skal kunne svare på forskningsspørsmålet «*Hvorfor opplever befolkningsrike land oftere utbrudd av intern konflikt enn andre?*». Hvert delkapittel i teorigjennomgangen starter med en grafisk fremstilling av korrelasjonen mellom aspektet som beskrives og befolkning. Denne korrelasjonen er en forutsetning for at de aktuelle variablene skal kunne sies å påvirke forholdet mellom befolkning og konflikt.

Deretter følger en gjennomgang av datasettet og de ulike variablene som er operasjonalisert utfra de teoretiske argumentene. En gjennomgang av oppgavens metodegrunnlag og ulike valg som er gjort hva gjelder datagrunnlaget, er også inkludert. Her er det blant annet inkludert en korrelasjonsmatrise for alle de inkluderte variablene, og diverse tester for å sjekke for eventuelle svakheter i datasettet. Kapittel 4 inkluderer de ulike modellene som har blitt kjørt, sammen med kommentarer til disse. Først presenteres modellene som er kjørt på hele datagrunnlaget. Dette er hovedresultatene, og analysen vil i stor grad bygges opp rundt disse resultatene. Deretter presenteres modellene som er kjørt på datagrunnlaget der Kina og India er ekskludert. En såkalt «fullmodell» kommenteres også. Til slutt i dette kapittelet sammenfattes resultatene fra de ulike modellene, og hypotesene vil enten forkastes eller beholdes avhengig av hvilke føringer de ulike modellene legger for hver enkelt hypotese.

Avslutningsvis vil oppgaven sammenfattes, og peke fremover på hva det trengs å undersøke nærmere for å få en klarere forståelse på hvorfor befolkningsrike land oftere opplever intern konflikt enn mindre befolkede land. Av oppgavens hypoteser er det særlig hypotesene som omhandler areal, etnisk relevante og etnisk ekskluderte grupper som viser seg å gi robuste resultater. Det foreslås derfor blant annet å gjøre et mer detaljert arbeid inn mot de spesifikke aspektene for å få en bredere forståelse for hvordan de ulike aspektene påvirker det eksisterende forholdet mellom befolkning og konflikt.

2. Teorigjennomgang

Størrelse, når det er mål i befolkningstall, spiller en viktig rolle for et lands plass og opptreden i verden, og er av stor interesse, blant annet på grunn av at majoriteten av verdens befolkning bor der. Samtidig opplever de også oftere intern konflikt. Selv om disse landene har til felles at de er folkerike, kan det være andre aspekter som også knytter dem sammen, og som kan forklare hvorfor befolkningsstørrelse påvirker intern konflikt. I dette kapittelet vil de eksisterende teoretiske argumentene for at ulike aspekter ved land påvirker sannsynligheten for intern konflikt gjennomgås, sammen med en redegjørelse for hvordan disse aspektene kan knyttes opp mot befolkningsstørrelse.

De ti mest folkerike landene i verden per dags dato, oppgitt fra størst til minst, er Kina, India, USA, Indonesia, Brasil, Pakistan, Nigeria, Bangladesh, Russland og Mexico (FN-sambandet, 2020). Ifølge Uppsalas konfliktdata har samtlige av disse landene hatt én eller flere interne konflikter i løpet av de siste ti årene, med unntak av Mexico og USA (UCDP, udatert). Selv om Mexico ikke har interne konflikter mellom staten og opprørsgrupper, har de mye konflikt i henhold til andre konflikttyper i Uppsalas datasett, da de blant annet opplever en betydelig mengde ikke-statlig konflikt mellom bander og gjenger (UCDP, udatert). Til sammenlikning med de ti minst befolkede landene (men med over 100 000 innbyggere), ser man tydelig at det er mer konflikt i de mer befolkede landene. Dette visualiseres i Tabell 1 og 2.

Land med størst befolkning i 2018	Antall kamprelaterede dødsfall 2008-2018
Kina	49
India	7930
USA	0
Brasil	440
Pakistan	37687
Nigeria	17031
Bangladesh	103
Russland	3130
Mexico	0

Tabell 1: Oversikt over konfliktrelaterede dødsfall i de mest befolkede landene i 2018 (Data fra FN og UCDP)

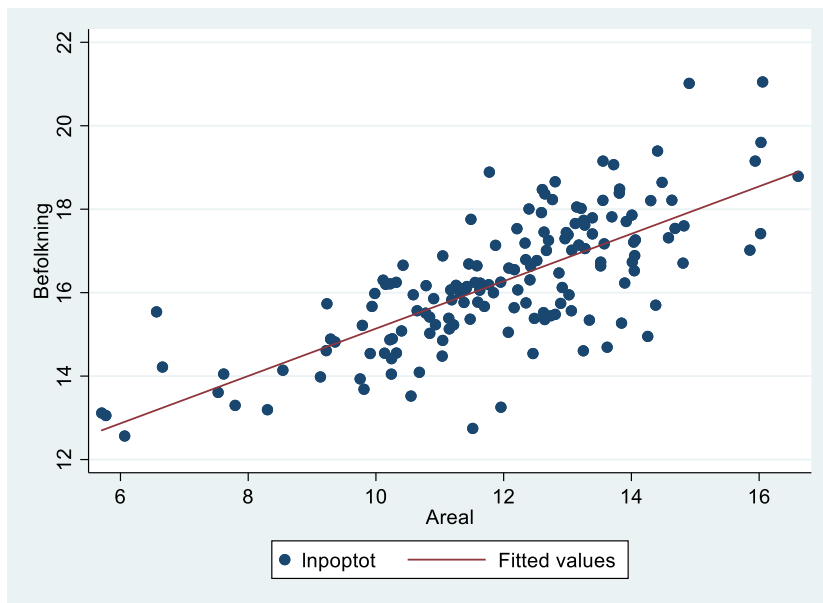
Land med mist befolkning (over 100 000) i 2018	Antall kampfrelaterede dødsfall 2008-2018
Barbados	0
Vanuatu	0
Sao Tome og Principe	0
Samoa	0
Saint Lucia	0
Kiribati	0
Saint Vincent og Grenadinene	0
Tonga	0
Grenada	0
Antigua og Bermuda	0

Tabell 2: Oversikt over konfliktrelaterede dødsfall i de minst befolkede landene i 2018 (Data fra FN og UCDP)

I kommende delkapitler vil fire ulike aspekter som alle karakteriserer verdens land, gjennomgås som mulige forklaringer på hvorfor store land oftere opplever borgerkrig enn små. Disse aspektene er plukket ut fordi de ifølge litteraturen gir større sannsynlighet for konflikt, og fordi de korrelerer med befolkningsstørrelse, som vil si at høye/lave verdier på den ene variabelen sammenfatter med høye/lave verdier på den andre variabelen. Samtidig er de plukket ut for at det så godt som mulig skal være aspekter som skiller seg fra hverandre, og ikke korrelerer for mye med hverandre. Disse aspektene er: (1) geografisk størrelse, (2) heterogen befolkning, (3) større masse å rekruttere fra, og (4) økonomiske fordeler ved å styre landet.

2.1 Geografisk størrelse og avstand

Et land som har stor befolkning vil også potensielt kunne ha et stort areal, og dermed vil utfordringene knyttet til store områder muligens kunne forklare hvorfor befolkningsstørrelse har noe å si på konfliktrisiko. Figur 1 nedenfor viser korrelasjonen mellom befolkningsstørrelse og størrelse målt i areal for alle landene som er med i analysen i 2017. I metodekapittelet vil korrelasjonskoeffisientene diskuteres nærmere, men prikkdiagrammet presenteres allerede her for å underbygge påstanden om at befolkningsstørrelse og et lands totale areal korrelerer med hverandre. Som en kan se i Figur 1, er det en tydelig positiv sammenheng mellom størrelse på befolkning og landareal.



Figur 1: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert totalt areal i 2017.

Ettersom befolkningsstørrelse og geografisk størrelse korrelerer slik de gjør, kan det altså tenkes at geografien til befolkningsrike land er med på å forklare sammenhengen mellom befolkningsstørrelse og utbrudd av konflikt. Geografi er blant annet en fundamentalt viktig variabel for å forstå ledelsen, overblikket og kontrollen i en opprørsgruppe samtidig som det kan fortelle viktige ting om statens sannsynlige styrke. I tillegg involverer en borgerkrig alltid to militære organisasjoner som kjemper om kontroll over et territorium, enten i form av en opprørsgruppe som søker kontroll over staten, eller som ønsker å separere en part av territoriet for å bygge en egen, ny stat (Buhaug & Rød, 2006; Raleigh & Hegre, 2009). Territorium kan også karakterisere borgerkrig gjennom blant annet avstand til viktige holdepunkter innad i landet, slik som hovedstat og opprøreres hovedkvarter (Toft, 2014:191). Denne avstanden til, og mellom, de respektive holdepunktene, påvirker forståelsen av egen relativ makt og styrke, som igjen påvirker en gruppes sannsynlighet for å gå i konflikt med andre. Et geografisk område har altså en sentral plass i en konflikt, uansett hvordan man ser på det.

Et lands areal kan også påvirke hvor det er sannsynlig at konflikter utspiller seg. Væpnede konflikter er forventet å skje i områder der makten til begge parter er relativt lik, ettersom et asymmetrisk maktforhold trolig ikke vil føre til konflikt på grunn av partenes subjektive forståelse av hjemmestyrke sammenliknet med den andre partens styrke (Butler & Gates, 2009:330-331). En gruppe som selv mener at de vil tape et opprør, vil ikke starte det i utgangspunktet. Statens styrke vil derfor ha mye å si i et slikt perspektiv. Byråkratiske og

økonomisk utviklede land vil for eksempel virke immune for konflikt nær hovedstaden, fordi de er så sterke at opprørsgruppene anser sin styrke til å være så liten i forhold til statens, at sannsynligheten for å vinne opprøret er minimal (Buhaug, 2010; Butler & Gates, 2009). Det betyr nødvendigvis ikke at konflikten alltid oppstår lengst unna hovedstaden, ettersom staten kan anslå maktforholdet til å være i favør av opprørsgruppen, og dermed anse det som lite sannsynlig at de kommer til å vinne en kamp i et slikt område. Dette kan blant annet være på grunn av vanskelig terreng, skog- og fjellområder (Carter m.fl., 2019:1448). Konflikter vil mest sannsynlig utspille seg i områdene der begge gruppene er usikre på hvorvidt de er sterkere enn den andre parten, noe som ofte, men ikke alltid, er et stykke unna hovedstaden.

Bouldings (1962) «Loss-of-strength gradient» er nøkkelenkonseptet som linker sosial og fysisk geografi til konflikt i forskningen på mellomstatlig konflikt. Buhaug (2010) gjenbraker Kenneth Bouldings «Loss-of-strength-gradient» på intern konflikt, og dette utgjør et konkret eksempel på hvordan makt basert på avstander også kan påvirke intern konflikt, på lik linje som mellomstatlig konflikt. Hovedpoenget er at konflikter ikke vil utspille seg i områder der en av partene er mye sterkere enn den andre (Buhaug, 2010). To opprørsgrupper med lik styrke, men med ulik distanse til statens maktsentrum, vil mest sannsynlig ha forskjellig sjanse til å gjøre opprør på grunn av beliggenhet. Konflikt vil på så måte være en funksjon av partenes relative muligheter (Cunningham, Gleditsch & Saleyhan, 2009:571). Distansen fra populasjonssentrum og perifere lokasjoner for opprørsaktiviteter spiller derfor en viktig rolle i å anta potensialet for opprørsgrupper til å organisere seg, rekruttere og trene (Linke m.fl., 2017:522).

Collier og Hoeffler (2004) drøfter hvorfor disse områdene er mer utsatt for konflikt og hvorfor staten ofte har mindre mulighet til å opprettholde makt og kontroll jo lengre ut i periferien man kommer. De konkluderer med at det kan være på grunn av mangel på støtte fra befolkningen, eller vanskeligheter med å styre og kontrollere områder langt unna statens hovedstad. Statlig militære, politi og andre styrker har større sannsynlighet for å være svake i rurale og geografisk marginaliserte områder av et land (Pickering, 2012; Linke m.fl., 2017:522). Dette er fordi et lands nasjonale styrke er sterkest i hjemmebasen til landet, og synker jo lengre unna man kommer dette punktet.

Hvis befolkningen i tillegg har liten tillit eller tiltro til staten, vil det være en høyere terskel for å melde ifra om potensielle opprørsgrupper, samtidig som befolkningen også har større

insentiver for å delta selv. Dette gir muligheter for oppblomstring av opprørsgrupper i områdene langt unna hovedstaden og maktsentrumet (Butler & Gates, 2009:331). I områder nærmere hovedstaden skal det for eksempel mer til for at en opprørsgruppe klarer å vokse seg stor og sterk fordi det kan forventes sterkere statlig kontroll og overvåking. Fearon og Laitin (2003) viser til hvordan vold blir mer sannsynlig når staten ikke klarer å finne eller slå ned på utfordrere, på grunn av deres svakhet eller mangel på kapasitet.

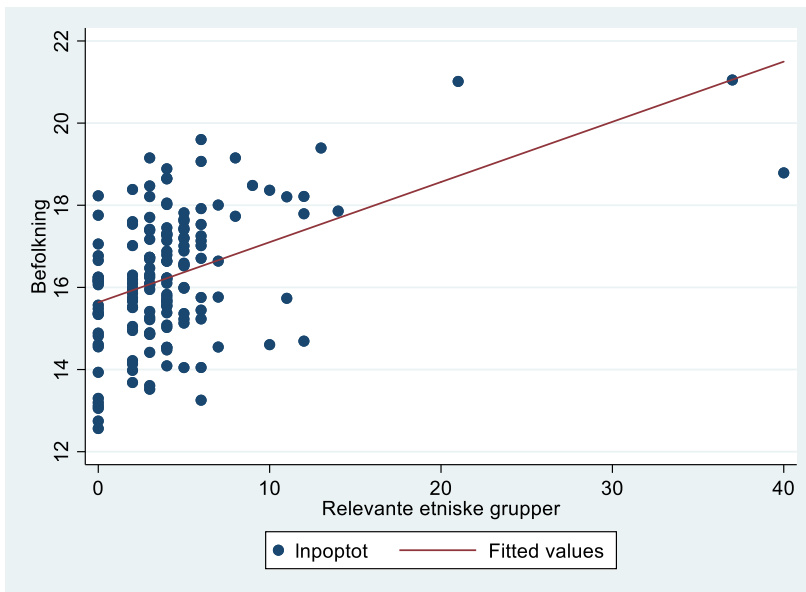
Siden land med stor befolkning ofte også er store i areal, slik som prikkdiagrammet i Figur 1 viser, kan man argumentere for at grunnen til at land med stor befolkning opplever mer borgerkrig enn små er som følge av ringvirkningene det totale arealet til staten gir. Et stort land vil ha større problemer med å styre landet og opprettholde ro, orden og kontroll langt unna hovedstaden, noe som åpner opp for potensielle opprørsgrupper. Et stort areal gir store avstander, som øker sannsynligheten for konflikt. Ut fra overstående betraktninger dannes hypotese 1:

H1) Land med stor befolkning opplever mer borgerkrig fordi de også har store territorium som kan gjøre det utfordrende for staten å ha kontroll på hele området sitt, og som dermed gjør det lettere for befolkningen å gjøre opprør.

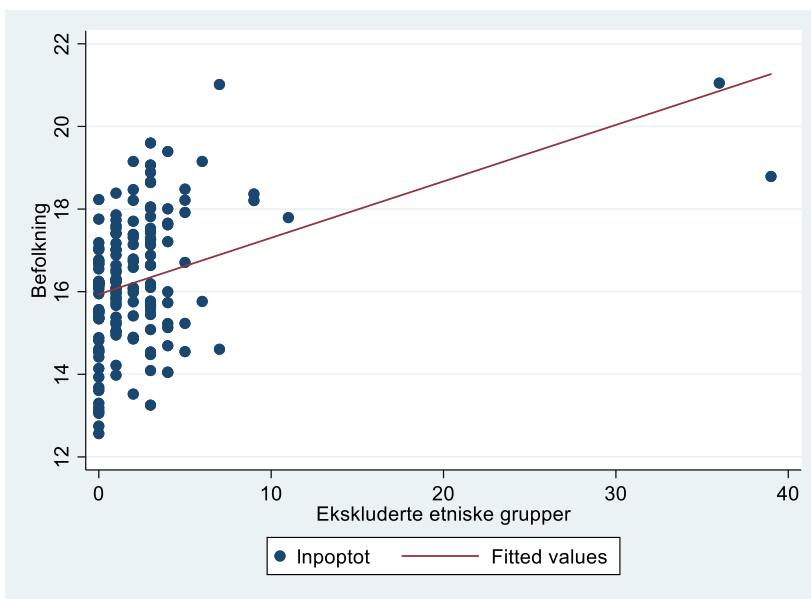
2.2 En heterogen befolkningsgruppe

Man kan anta at det er større sannsynlighet for at to tilfeldig valgte individer ikke tilhører den samme gruppen i samfunnet, enten det er etnisk, religiøst eller språklig, i folkerike land. I et land med mange mennesker, vil det være sannsynlig at det finnes flere ulike grupper enn i mindre land, og dermed også flere potensielle grupper som er ekskludert fra det politiske liv. Prikkediagrammene i Figur 2 og 3, viser hvordan befolkningsstørrelse og antall relevante etniske grupper, samt befolkningsstørrelse og antall ekskluderte etniske grupper korrelerer med hverandre i 2017. I metodekapittelet vil korrelasjonskoeffisientene diskuteres nærmere, men prikkdiagrammet presenteres allerede her for å vise eventuelle sammenhenger mellom variablene. Som en kan se henger stor befolkningsstørrelse sammen med flere grupper i begge

figurene.



Figur 2: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og antall relevante etniske grupper i 2017.



Figur 3: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og antall ekskluderte etniske grupper i 2017.

Ettersom befolkningsstørrelse og antall etniske grupper som er relevante eller ekskluderte korrelerer slik de gjør, kan det altså tenkes at mangfoldet i befolkningsrike land er med på å forklare sammenhengen mellom befolkningsstørrelse og utbrudd av konflikt. Det er kun et fåtall av alle landene i verden som er homogene, som vil si at de fleste land består av flere enn én gruppe mennesker. De heterogene landene kan imidlertid ha stor forskjell i grad av mangfold, der noen består av veldig mange grupper, mens andre kun består av et par. India

har for eksempel 21 relevante etniske grupper i 2017, mens Island er ett av få land som kun er oppført med 1 relevant etnisk gruppe.

Under den kalde krigen utviklet det seg en tanke om at etniske identiteter skulle bli byttet ut mot lojalitet til et større samfunn, og at man dermed ville oppleve færre konflikter over etniske skillelinjer (Gurr, 1994:347; Cederman m.fl., 2017:262). Det har tidligere vært en gjennomgående tanke at økonomisk utvikling, urbanisering og en økende andel skrivekytting skulle føre til skapelsen av komplekse og integrerte samfunn gjennom hele verden.

Moderniseringsteorier, som samfunnsvitere aksepterte mer eller mindre, predikerte at mer politisk og økonomisk interaksjon mellom mennesker, og veksten av kommunikasjonsnettverk, skulle gjøre folk mer lojale til større samfunn slik som EU og «Pan-Afrika», og la etnisk opphav bli erstattet av en slik lojalitet (Gurr, 1994:347-348). Likevel er et stort antall av dagens borgerkriger relatert til identitet, og er ofte kjempet mellom ulike etniske eller religiøse grupper. Som en konsekvens av dette, er det mange som antar at etnisk og religiøs dynamikk kan gjøre sannsynligheten for borgerkrig større (Buhaug & Gates, 2002:420).

Forskning på politisk konflikt har utpekt tre forhold som påvirker om en gruppe mobiliserer og utøver vold. Disse er felles identitet, frustrasjon og mulighet (Ellingsen, 2000:229). Disse tre forholdene kan styrkes hvis etnisitet spiller en betydelig rolle for mobiliseringen. Alle individer identifiserer seg med noe, enten det er kjønn, generasjon, organisasjon, sosial klasse, etnisk gruppe eller nasjonalstat. Ofte identifiserer vi oss med flere slike subgrupper, avhengig av konteksten (Smith, 1993). Det som gjør etnisk identitet spesielt utsatt for konflikt sammenliknet med andre identiteter, er at identiteten er basert på fundamentale faktorer som språk, historie, etnisk gruppe eller religion. Disse faktorene, som henger sammen med familie og blodsband, er aspekter som folk ofte er villige til å kjempe og dø for. Etnisitet skiller seg i så måte fra andre typer identitet et individ kan ta ved at de er satt og ikke-diskuterbare (Reynal-Querol, 2002:29). Dette er faktorer som ofte virker viktigere enn territorielle grenser, og som sjeldent matcher disse grensene (Ellingsen, 2000:229; Gurr, 1994:348; Reynal-Querol, 2002:29). Dette er et primordialistisk syn på etnisitet, mens andre, som blant annet konstruktivister og instrumentalister, vil hevde at etnisitet og identitet kan formes og endres. Det finnes altså flere motstridende syn på hvordan etnisitet påvirker mennesker.

Lake og Rotchild (1998:7) argumenterer for at etnisitet i seg selv ikke er en årsak til konflikt, men når etnisitet henger sammen med opplevd sosial ulikhet, konflikthistorie og frykt for hva fremtiden bringer, er etnisitet en av linjene samfunnet fordeler seg utfra. Om opplevd sosial ulikhet også oppleves å være på grunn av egen identitet, vil mennesker som føler på den samme frustrasjonen samle seg som gruppe. Det er altså ikke etniske uenigheter i seg selv som driver konfliktene, da de er drevet av den samme frustrasjonen eller misnøyen som ikke-etniske konflikter drives av (Denny & Walter, 2014:199). På denne måten kan etnisk fraksjonalisering ha en effekt på sannsynligheten for intern konflikt, men uten at den er direkte, og nødvendigvis synlig (Blimes, 2006:539).

En opprørsgruppes mulighet til å kunne gjøre opprør, avhenger av både rekrutteringsmuligheter, finansiering og statlig styrke (Danny & Walter, 2014:204). Et felles språk og en felles kultur gjør det lettere for etniske opprørsledere å spre informasjon for å øke muligheten til å rekruttere flere individer (Bormann m.fl., 2017:748). Nærhet i form av liten distanse er også en viktig faktor som spiller inn (Gates, 2002). I tillegg vil ofte båndene mellom medlemmene i en opprørsgruppe med individer fra samme etniske gruppe være sterkere enn i en opprørsgruppe som er samlet på bakgrunn av andre faktorer (Danny & Walter, 2014:204). Etniske grupper har også større sannsynlighet for å få økonomisk støtte utenfra, og dermed finansiere et mulig opprør. Siden migranter ofte bosetter seg blant andre migranter fra samme land, oppstår det store diasporaer som har sterke bånd til hjemlandet sitt, og derfor også er villige til å bidra med økonomisk støtte (Danny & Walter, 2014:204).

Bare et fåtall av verdens land er som sagt språklig, religiøs og etnisk homogene. Etnisitet handler om faktorer som delt historie, tro, religion, språk eller hjemland, og det gjør ofte konflikter mer gjennomgripende (Ellingsen, 2000:229). Etnisk komposisjon operasjonaliseres vanligvis over to dimensjoner i litteraturen. Den første dimensjonen er fragmentering; jo flere grupper, eller jo høyere sannsynlighet for at to tilfeldig valgte individer hører til ulike grupper, jo høyere nivå av fragmentering i en stat. Den andre dimensjonen er polarisering eller dominans (Cunningham & Weidmann, 2010:1038). I denne oppgaven vil det først og fremst være fraksjonaliseringsaspektet som operasjonaliseres og brukes, ettersom det er antall grupper som antas å henge sammen med befolkningsstørrelsen til et land.

I fragmenterte samfunn kan det være vanskelig å definere hva «folket» er, og med denne utfordringen følger også gjerne sannsynligheten for konflikt (Esteban & Ray, 1999; Elbadawi

& Sambanis, 2002; Collier & Hoeffler, 1998). Litteraturen på dette feltet har ofte tatt utgangspunkt i «etnolingvistisk fraksjonaliseringsindeks» (ELF), som måler sannsynligheten for at to tilfeldige mennesker tilhører forskjellige grupper. En høyere ELF vil da gå sammen med høyere risiko for konflikt. Likevel argumenteres det også for at politikken i land med veldig heterogen befolkning vil kunne ligne strategisk på politikken i veldig homogene land, siden de små etniske gruppene må skape koalisjoner for å ha politisk påvirkning (Horowitz, 1985; Fearon & Laitin, 2003:6). Politikken i disse landene er da preget av samarbeid, og de fleste gruppene blir inkludert i de politiske avgjørelsene. Derfor antar noen at antall grupper innad i et land er relatert til borgerkrig i en omvendt U-kurve, der antall konflikter er høyest der det er flere grupper, og mindre der det er få eller veldig mange grupper (Ellingsen, 2000:233). En slik tankegang går ikke sammen med tanken om at jo mer fraksjonalisert en befolkning er, jo større sannsynlighet er det for konflikt. Det er altså uenigheter i om hvorvidt et høyt antall grupper skaper konflikt eller ikke.

På den andre siden er det andre som mener at det ikke er antall grupper som øker sannsynligheten for konflikt, men hvor mange av disse gruppene som er marginaliserte eller ekskluderte i samfunnet de lever i (Easterly & Levine, 1997; Cunningham & Weidmann, 2010:1035). De diskriminerte eller ekskluderte gruppene blir fremmedgjort i det de anser som sin egen stat, ettersom staten styres av «fremmede» individer som tilhører andre etniske grupper enn dem selv (Cederman m.fl., 2013:59; Gurr, 1993:161). Hvis en etnisk gruppe sitter med mye makt i regjering, kan de høste goder i form av statsborgerskap, muligheter til jobber i staten og så videre, mens grupper som er totalt ekskludert fra regjeringen, eller ikke er representert, går glipp av disse godene (Bormann m.fl., 2017:748; Cederman m.fl., 2013:61). Denne etnopolitiske favoriseringen fører med seg et politisk system der borgere med samme etnisk gruppe, eller som er bosatt i samme område som politikere med mye makt, høster klare fordeler sammenliknet med andre (Theisen, Strand & Østby, 2020:4-5). Dette kan være grobunn til frustrasjon, misnøye og opplevelse av urettferdighet, som igjen kan føre til mobilisering og konflikt (Fjelde & Østby, 2014).

Siden staten sjeldent er etnisk nøytral, kan den være med på å utvikle, og sette i gang, etniske konflikter (Cederman, Wimmer & Min, 2010:95; Cunningham & Weidmann, 2010:1040). Selv i multietniske land vil en gruppe som oftest være mer assosiert med staten enn de andre, og staten i seg selv vil være den ultimate pris for enhver etnisk gruppe å vinne fordi den bærer med seg store fordeler (Bormann m.fl., 2017:753; Cederman m.fl., 2013:61). Der grupper er

marginalisert, politisk ekskluderte eller er redde for å miste muligheten til å ivareta sin etniske identitet, kan det oppstå frustrasjon som kan føre til konflikt (Denny & Walter, 2014:202).

Basert på tidligere forskning kan man altså si at det ikke finnes et enkelt og klart svar på hvorvidt etnisk mangfold har noen påvirkning på borgerkrig, og hvordan denne påvirkningen eventuelt skjer. Likevel ønsker jeg å trekke frem argumentet om at jo flere grupper det finnes innad i et samfunn, jo større er sannsynligheten for at noen av de vil være marginalisert eller ekskludert, ettersom dette kan påvirkes av befolkningsstørrelse. Tidligere forskning har pekt på at både flere etniske grupper og etnisk ekskluderte grupper, kan føre til større sannsynlighet for borgerkrig. I denne studien vil det bli testet om disse faktorene også kan være med på å forklare hvorfor land med stor befolkning opplever mer borgerkrig enn andre. Med bakgrunn i dette dannes hypotese to og tre:

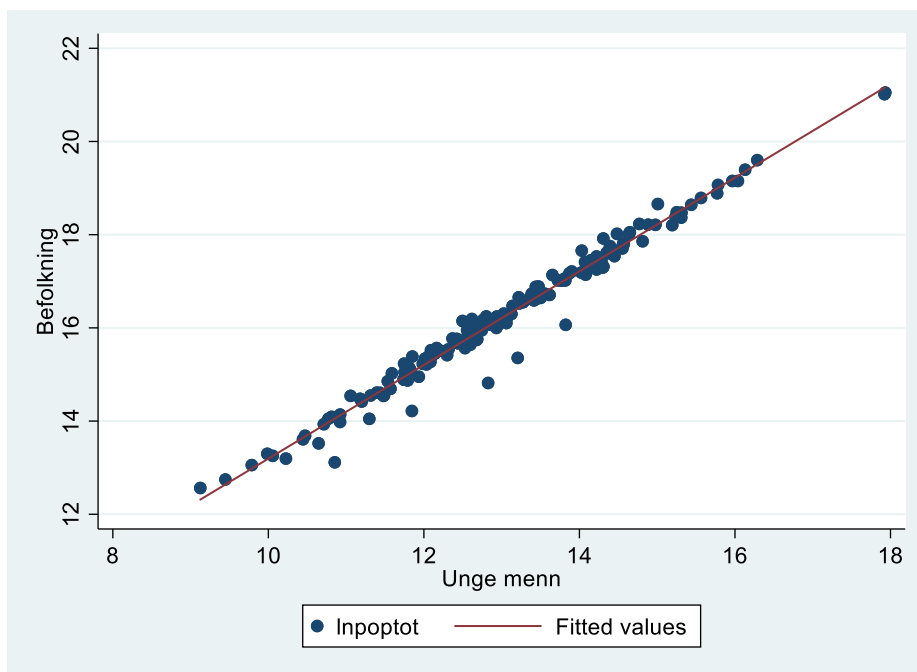
H2: Land med stor befolkning opplever mer borgerkrig fordi det er flere etniske grupper innad i befolkningen som kan være misfornøyde og gå til opprør

H3: Land med stor befolkning opplever mer borgerkrig fordi det er flere ekskluderte etniske grupper innad i befolkningen som kan være misfornøyde og gå til opprør

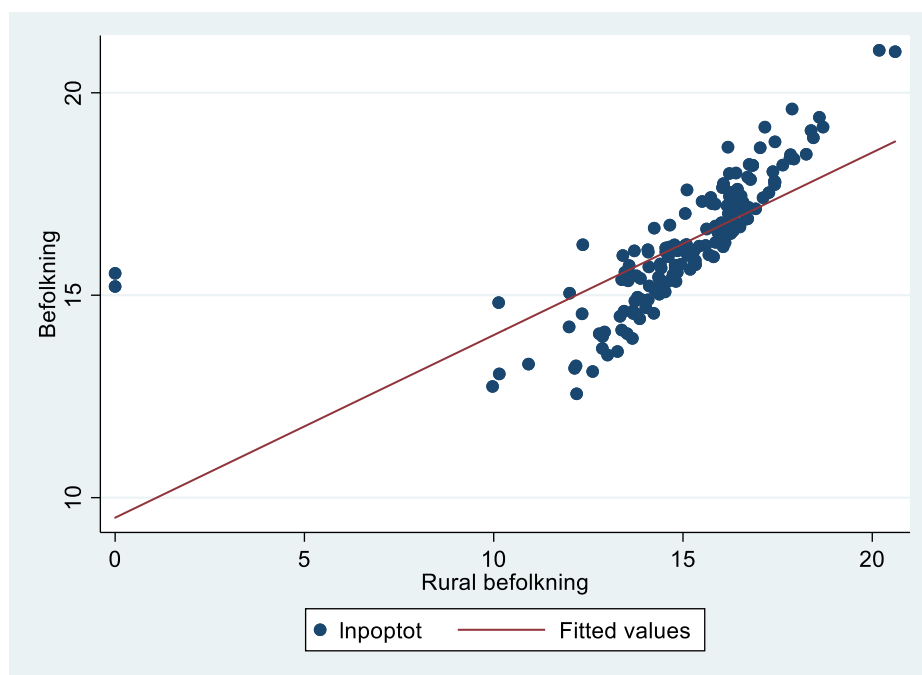
2.3 En større masse av potensielle opprørere

Man kan anta at et land med stor populasjon vil kunne bestå av et større absolutt antall individer som enten er marginalisert på bakgrunn av identitet eller samfunnsklasse, eller som er fattige og derfor har lite å tape ved å søke et bedre liv gjennom opprør. Et større absolutt antall individer som passer inn i beskrivelsene av hva som kjennetegner en opprører fra en ikke-opprører, gjør også at opprørsgruppene til enhver tid vil ha en større pool å rekruttere fra. Slik kan befolkningsstørrelse henge sammen med opprørsgruppenes muligheter til å vokse seg store og betydningsfulle.

Prikke-diagrammene i Figur 4 og 5 viser hvorvidt befolkningsstørrelse og antall unge menn, og befolkningsstørrelse og størrelse på den rurale befolkningen i landet, korrelerer med hverandre. I metodekapittelet vil korrelasjonskoeffisientene diskuteres nærmere, men prikke-diagrammene presenteres allerede her for å vise den eventuelle sammenhengen mellom variablene. Som en kan se er det en positiv sammenheng mellom befolkning og rural befolkning og befolkning og unge menn.



Figur 4: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert unge menn i 2017.



Figur 5: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert rural befolkning i 2017.

Når det er lettere for opprørsledere å rekruttere nye opprørere, vil det være større sannsynlighet for at flere grupper har mulighet til å bli store og betydningsfulle. Selv om det finnes mye litteratur om emnet, er det fortsatt debatt om hvilke omstendigheter som motiverer kvinner og menn til å delta i dødelige kamper (Humphreys & Weinstein, 2008:436;

Cederman, Wimmer & Min, 2010). Motivasjonen for å delta i opprør kan komme fra individets frustrasjon, eller fra ønsket om å handle på vegne av sin egen sosiale eller økonomiske gruppe (Humphreys & Weinstein, 2008:439). Olsons (1965) analyse av kollektiv handling gir oppstand til to tilnærminger som setter søkelys på at individene veier kostnadene opp mot gevinstene når de vurderer om de skal delta i et opprør. Den første tilnærmingen vektlegger viktigheten av at deltakelse må være til fordel ikke bare for gruppen som helhet, men også for individene selv. Det kan blant annet bety at det kreves private goder i bytte mot deltakelse. Godene kan for eksempel være i form av penger, landområder eller viktige posisjoner i staten. Andre faktorer vil også kunne påvirke valget til individene om hvorvidt de skal mobilisere eller ikke, slik som dedikasjon til saken, risikotaking eller holdninger til vold (Berman m.fl., 2011:501). Det er altså av oppfatning at misnøye og vurderingen om kostnader opp mot fordelene ved å delta er viktige faktorer som forklarer deltakelse i opprørsgrupper. Disse faktorene er såkalte «pull-faktorer», som drar individer mot å bli opprørere.

Kritikere som mener at denne fortolkningen av Olson er for materialistisk, fokuserer heller på viktigheten av sosiale sanksjoner. De hevder at samfunnet kan utøve et sosialt press på individet som endrer hvordan kostnader og fordeler evalueres opp mot hverandre (Humphreys & Weinstein, 2008:439; Gates, 2002:116). Taylor (1988) definerer et sterkt samfunn som et medlemskap med delte verdier og tanker, med relasjoner innad i samfunnet som er direkte og mangesidige, og med praktiseringer innenfor samfunnet som er generalisert gjensidig. Sosiale sanksjoner kan virke som «push-faktorer», der det å delta i et opprør kan oppfattes som eneste mulighet. Det kan for eksempel komme i form av trusler om at individene blir utsatt for vold hvis de ikke selv deltar i opprørsgruppen, eller at de setter familie og venner i fare (Humphreys & Weinstein, 2008:441; Gates, 2002:116).

Fattige mennesker vil ofte ha større insentiver til å delta i opprør enn andre grupper, fordi de har mindre å tape, og mer å vinne. Mulighetskostnader og predikasjoner om at voldelig opprør vil øke når ulikhetene i inntekt stiger er en kollektiv oppfatning, ettersom de relativt fattige opprørerne vil se at de kan tjene mer på å gjøre opprør mot de rike enn å beholde status quo (Berman m.fl., 2011:498). Studier viser også at antall kriminelle handlinger øker når bølger med økonomisk nedgang slår inn, og arbeidsledigheten stiger (Berman m.fl., 2011:499). Dette vises også i store kvantitative undersøkelser som finner at en lav BNP per capita predikerer borgerkriger (Fearon & Laitin, 2003).

Mennesker som av en eller annen grunn er misfornøyde eller fattige relativt til resten av befolkningen, kan altså ha en større sannsynlighet for å delta i voldelige opprør. Folk som bor et stykke unna sentrum i et land kan høre til en slik gruppe. Siden befolkningen som bor i rurale områder er lengre unna hovedstaten der de offentlige godene blir produsert, får de ofte færre goder for skattepengene sine sammenliknet med de andre (Buhaug, 2006:697). Denne opplevde urettferdigheten kan føre til opprør. Alesina & Spolaore (2003) argumenterer også for at statlig størrelse er positivt assosiert med heterogenitet i preferanser i befolkningen. Jo flere mennesker, og eventuelt jo større avstander mellom dem, jo flere ulike preferanser vil det være. Dette gjør at det vil være vanskelig for staten å imøtekomme alles behov, og da er det naturlig at den rurale befolkningen blir nedprioritert til fordel for den urbane befolkningen. Den rurale befolkningen vil også ofte bestå av fattige mennesker sammenliknet med de som bor i større byer. Ser man på sammenhengen mellom BNP per capita og størrelse på befolkningen som bor i rurale områder er denne negativ. Det vil si at man generelt kan beskrive land med en stor rural befolkning som fattigere og med flere ulike preferanser enn land med mindre rural befolkning. Den rurale befolkningen befinner seg også i områder der staten er svak, som i tillegg gjør sannsynligheten for opprør større (Herbst, 2002; Sommers, 2003).

En annen gruppe som kan være mer disponert for å gjøre opprør enn andre, er unge menn. Unge mennesker står ofte i front når det oppstår revolusjoner, og det er ofte unge menn som står for mesteparten av de voldelige handlingene i en konflikt (Fluckiger & Ludwig, 2018:1937; Beber & Blattman, 2013). En stor ungdomsgruppe kan utvikle en bevissthet om sin egen størrelse og styrke, som motiverer dem til kollektiv handling (Urdal, 2011). Et stort antall unge menn i et land er dermed antatt å gjøre sannsynligheten for borgerkrig større (Yair & Miodownik, 2016:25; Fluckiger & Ludwig, 2018:1932). Dette er blant annet fordi en økning i proporsjonal størrelse på unge mennesker i et land minker statens kapasitet, særlig i form av utdanningsmuligheter og arbeid. Mange unge mennesker kan føre til økt arbeidsledighet og lavere økonomisk utfall enn forventet for ungdomsgruppene (Yair & Miodownik, 2016:26). En økt arbeidsledighet og lavere lønninger kan igjen føre til at flere unge ser det verdifullt å delta i opprør for å endre status quo, ettersom de har lite å tape (Urdal, 2011). I tillegg vil unge menn oftere bare ha seg selv å tenke på, til forskjell fra de som er eldre og har familier å ta vare på. Mange mener blant annet at den arabiske våren var et resultat av mange unge voksne som hadde lite å tape på å utfordre normaltilstanden. I gjennomsnitt var 42% av befolkningen i Nord-Afrika og Midtøsten mellom 15 og 29 år i

2010, mens denne gruppen utgjorde 35% i resten av verden dette året (Yair & Miodownik, 2016:25).

Basert på dette kan man anta at det i land med en stor befolkning vil finnes et større absolutt antall mennesker som er villige til å delta i voldelige opprør. Det kan blant annet være fordi de har andre preferanser enn hovedstadsområdet i landet, eller fordi de for eksempel er unge og ugifte, og dermed ikke har noen å ta vare på hjemme. Felles for disse gruppene er at de ofte er fattigere enn majoriteten, eller opplever urettferdig fordeling av goder. Fordi det finnes et fåtall av gode mål på ulikhet som er til å stole på, og som i tillegg har kontinuerlige målinger, vil fattigdomshypotesene operasjonaliseres med andre mål enn de typiske målene som GINI, BNP per capita og absolutt/relativ fattigdom. BNP per capita har blitt forsøkt brukt, men er ikke et godt mål på fattigdom fordi det bare er et mål på gjennomsnittlig inntekt i et land, i tillegg til at det korrelerer dårlig med befolkningsstørrelse. Antall unge menn og antall mennesker som bor i rurale områder er derfor valgt ut til å representere fattigdomsaspektet i analysen. Basert på dette dannes hypotese fire og fem:

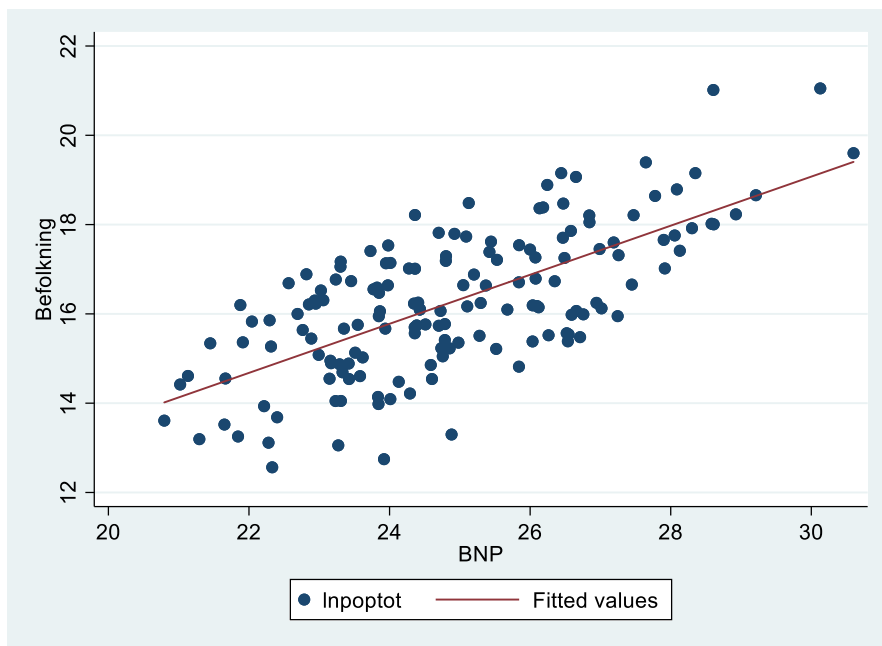
H4: Land med stor befolkning opplever mer borgerkrig fordi det består av flere unge menn som anser det som mer lønnsomt å delta i voldelig konflikt enn å bevare status quo.

H5: Land med stor befolkning opplever mer borgerkrig fordi det består av flere mennesker som bor i rurale områder som anser det som mer lønnsomt å delta i voldelig konflikt enn å bevare status quo.

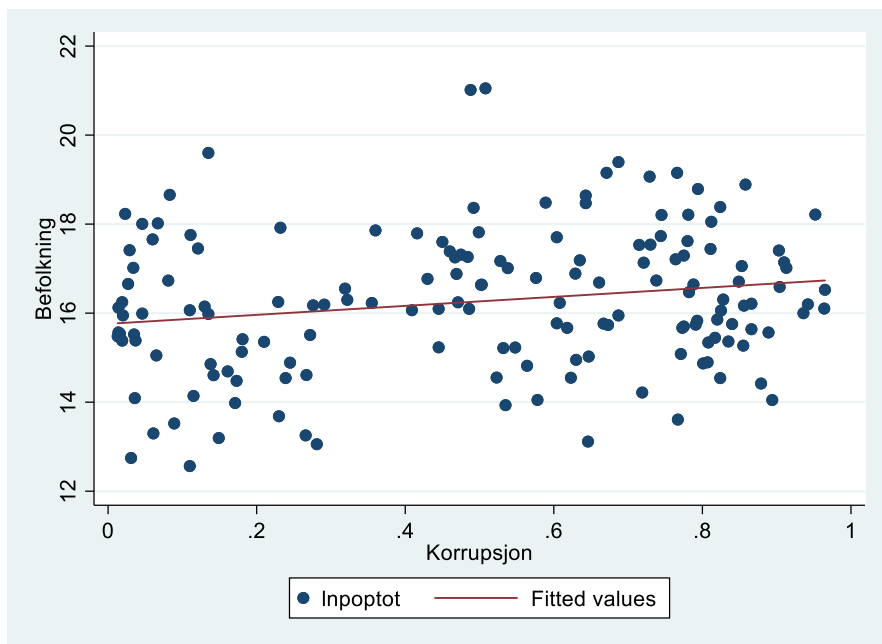
2.4 Store gevinster å hente ved å styre

I et stort land med mange mennesker vil det som oftest finnes mye ressurser. Mange sier at mennesket er den største ressursen av dem alle, spesielt med tanke på arbeidskraft og intellekt (se f.eks Manuelli & Seshadri, 2014; Bravo-Ortega & de Gregorio, 2005). Denne tanken baserer seg på at mennesket bidrar med arbeid, og tjener penger som staten kan skattlegge i flere år, som igjen gir staten inntekter. Når staten har store muligheter til å tjene penger, vil det også være interessant for enkeltindivider å sitte med makten, fordi det muligens kan bety verdier til dem. En stat med store inntekter og økonomi, kan også være mer utsatt for korrupsjon, som gjør det mulig å ta ut disse pengene til fordel for seg selv.

I prikkediagrammene i Figur 6 og 7, vises korrelasjonen mellom befolkningsstørrelse og totalt bruttonasjonalt produkt (BNP), og befolkningsstørrelse og korrupsjon. I metodekapittelet vil korrelasjonskoeffisientene diskuteres nærmere, men prikkediagrammet presenteres allerede her for å vise eventuelle sammenhenger mellom variablene. Som man kan se er sammenhengen mellom befolkning og BNP sterk, mens den ikke er like sterk mellom befolkning og korrupsjon.



Figur 6: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og logtransformert total BNP i 2017.



Figur 7: Grafisk fremstilling av korrelasjonen mellom logtransformert befolkningsstørrelse og korrupsjon i 2017.

Det finnes endel tidligere forskning som baserer seg på at opprørere er motivert av grådighet, og at man vil gjøre opprør hvis det er mer å vinne enn å tape på å mobilisere (Collier & Sambanis, 2002:4; Collier & Hoeffler, 2004). Jo større økonomisk gevinst det er å hente ved å kontrollere et område, jo større arbeid og ressurser vil opprørerne legge ned for å vinne og ha kontroll over det gitte området (Berman m.fl., 2011:498). Et område kan være verdifullt økonomisk av mange grunner, men ofte har det med tilgang på ressurser å gjøre. I tidligere forskning har mekanismene som knytter ressursrikdom til borgerkrig vært knyttet til politiske feiltrinn. Profitten som kommer fra naturressursene gjør at lederne vil underinvestere i statens infrastrukturelle styrke, da tilgang til store overskudd gjør lederne mindre avhengige av samfunnsmessige statsapparater, som fører til at de offentlige institusjonene blir svakere. Svake institusjoner for sosial kontroll gjør regjeringens mulighet til å overvåke og eventuelt stoppe opprørere før det blir voldelig svært hemmet, som igjen øker muligheten for opprør (Fjelde, 2009:201).

I land som er ikke-demokratiske, eller som sliter med mye korrupsjon, vil ofte eliten ha brukt statens økonomiske ressurser til egen nytte, som for eksempel for å beholde statsmakten eller bevare egne verdier (Boix, 2003). I slike samfunn har de mindre privilegerte bare én valgmulighet for å få det bedre, og det er revolusjon. I mange demokratiske land er det motsatt. Der er de økonomiske ressursene mobile og bundet opp på en slik måte at de ikke kan skattlegges. Eksempler på dette kan være kunnskap, aksjer og andre finansverktøy som ikke er fysisk til stede (Boix, 2003). Når ressursene ikke kan skattlegges, vil det heller ikke være mulig for den politiske ledelsen å spise av lasset, og godene blir fordelt utover den generelle befolkningen i stedet for å havne hos noen få. Å sitte med makten i slike land vil dermed ikke kunne gi de samme økonomiske fordelene som i mindre demokratiske land med mye korrupsjon. Derfor er det også ofte mindre konflikt i demokratier enn i land med andre styreformer (Hegre m.fl., 2001). Boix (2003) kan dermed tolkes som at rike land både er, og ikke er, ekstra utsatt for intern konflikt, og at det er statens kapitalmobilitet og grad av korrupsjon som påvirker utslaget.

Korrupsjon er spredt utover ressursfattige så vel som ressursrike land i verden. Mange afrikanske land består for eksempel av dyp fattigdom hos den store majoriteten, samtidig som tilgang på verdifulle naturressurser som olje, gass og diamanter gjør en liten del av befolkningen svært rik (Mcferson, 2009:1529). Svaret på hvorfor det er slik, ligger i den ekstreme korrupsjonen i slike land, der eliten kan kjøpe seg kontroll og hindre politisk

deltakelse fra resten av befolkningen (McFerson, 2009). Et korrumpert samfunn er fanget i et institusjonelt anarki med vold (Theobald, 1990:130). Egeninteresse gjør at sosiale ulikheter blir mer intense, og at det blir en økt sosial fragmentering i samfunnet som igjen fordrer til konflikt (Le Billion, 2003). Muligheten lederne har i korrumperte stater til å spise av lasset, gjør også at andre grupper vil se at de eventuelle verdiene av å styre landet vil være mye større enn kostnaden ved å gå i konflikt, og det vil derfor være mer interessant for grupper å prøve å overta makten over territoriet (Le Billion, 2003:419). Hovedformålet til gruppene som går i konflikt med staten er da ofte egeninteresse og egen velstand, ettersom korrupsjonen gjør det mulig å putte pengene i egen lomme i stedet for å gi til resten av befolkningen (Le Billion, 2003:421).

Basert på tanken om at land som er korruperte og ressursrike opplever borgerkrig på grunn av at det er mye å tjene på å sitte med makten, kan det argumenteres for at en stor befolkning kan gi liknende økonomisk gevinst. Hvis et land har en stor befolkning, kan man også anta at staten er stor og tjener store summer på å skattlegge egen befolkning. Dette fordi en stor befolkning vil bestå av en stor gruppe i arbeidsdyktig alder. Om en stat eier store verdier eller at et stort antall av befolkningen er i arbeid, vil det også bety at det er flere å skattlegge, og dermed også mer penger. Hvis staten i tillegg lider av korrupsjon, vil det kunne motivere opprørsgrupper til å mobilisere fordi de vet at det ligger store verdier i å sitte å styre staten. Basert på en slik antakelse dannes den siste hypotesen, hypotese seks:

H6: Land med stor befolkning opplever mer borgerkrig fordi det er mer å tjene på å styre landet på grunn av korrupsjon og tilgang på kapital.

2.5 Alternative teoretiske grunnlag

Det har også vært aktuelt å ta med en hypotese som handler om befolkningstetthet og kamp om ressurser som eventuell forklaring på sammenhengen mellom befolkning og konflikt. Blant annet Homer-Dixon (1999), Gleick (1993), Hauge og Ellingsen (2001), Gizelis & Wooden (2010) og Bao og Fang (2007) finner signifikante resultater på at konfliktrisiko blir høyere når det er knapphet på ressurser, og da særlig tilgang på ferskvann. Tir og Diehl (2001), Hauge og Ellingsen (2001), De Soysa (2002) og Raleigh og Hegre (2009) argumenterer også for at befolkningstetthet øker sannsynligheten for konflikt. Tilgang på ferskvann og befolkningstetthet kunne derfor vært interessante aspekter å inkludere i analysen, ettersom det er logisk å tenke at disse påvirkes av befolkningsstørrelse. Pearsons

korrelasjonskoeffisient viser derimot at det er en svak korrelasjon mellom variablene som operasjonaliserer disse aspektene, med henholdsvis $-0,117$ og $0,1514$. I følge Mehmetoglu og Jakobsen (2017:39) er verdier på $0,1$ svake korrelasjoner. For å teste om jordbruk eventuelt kunne fungere som et mål på ressurstilgang har det også her blitt gjort en korrelasjonsanalyse, men også denne viser en svak korrelasjon med befolkningsstørrelse på $-0,157$. Mangel på korrelasjon med befolkning har ført til at denne delen har blitt utelatt.

3. Data og metode

I denne delen av oppgaven vil metode- og datagrunnlaget gjennomgås. De første delene vil handle om datasettet og operasjonalisering av de ulike hypotesene. Deretter følger metodedelene av oppgaven. Her vil det være korte og konkrete diskusjoner og redegjørelser av viktige aspekter ved oppgaven. Oppgaven bygger følgelig på kvantitativ metode ved bruk av paneldata, og er gjennomført i STATA 15.1.

3.1 Gjennomgang av data

Datasettet som brukes i analysen for å teste hypotesene består av 7123 observasjoner fra 1961 frem til 2017, og har observasjoner for hvert land-år. Flere av variablene har opprinnelig målinger som går lengre bak i tid, men på grunn av manglende eller sporadiske målinger på andre variabler frem til 1961, har disse blitt fjernet. Dermed vil funnene kun gjelde for konflikter etter 1961. Et par land har også blitt fjernet fra utvalget, ettersom det viste seg vanskelig å finne gode målinger på de ulike variablene over lengre tid. Det kan trolig begrunnes med at de ulike datasettene følger ulike standarder for hva som regnes som et selvstendig land og ikke, slik som Hong Kong og Taiwan, men også forskjell i hvordan de har håndtert land/grupper av land som har byttet navn og størrelse i løpet av perioden. Dette gjelder for eksempel Jugoslavia og Sovjetunionen, som i denne analysen har blitt behandlet som landene de er i dag gjennom hele perioden. Enkelte land, som Andorra og andre mindre øystater, har også blitt utelatt fra studien, da datagrunnlaget også her var mangelfullt.

3.2 Avhengig variabel(Y) – Utbrudd av konflikt

Variabelen som brukes for å måle konflikt i denne oppgaven, er hentet fra Onset datasett 19.1 fra UCDP/PRIO, og måler utbrudd av intern konflikt i perioden 1946 til 2017 (Gleditsch m.fl., 2002). Valget om å bruke utbrudd av konflikt og ikke andre aspekter ved konflikt, slik som varighet, antall konflikter eller antall døde, er tatt først og fremst fordi befolkningsstørrelse har gitt robuste resultater som er gjort mot utbrudd av konflikt, slik som i Sambanis (2004) og Hegre & Sambanis (2006) sine analyser. Datasettet som brukes måler om det er utbrudd av en ny konflikt i det inneværende året i hvert enkelt land. Variabelen er dikotom, og tar verdien 1 ved et nytt utbrudd, og verdien 0 hvis det ikke er noe nytt utbrudd. Det vil si at det finnes observasjoner for hvert land i alle år, uavhengig av om de har opplevd en intern konflikt. For noen av landene i datasettet er det flere observasjoner for hvert enkelt år, som indikerer at det er to ulike konflikter som utspiller seg på samme tid. Disse observasjonene behandles som én.

Datasettet som brukes, har flere ulike variabler for utbrudd, der det som skiller dem fra hverandre er hvor lenge det er siden en konflikt hadde sin siste observasjon før det regnes som en ny konflikt. I denne oppgaven brukes onset2-variabelen, som betyr at en konflikt regnes som ny hvis det er mer enn to år siden siste observasjon i forrige konflikt (Pettersen, 2019). Det vil si at datasettet som brukes i analysen ikke forteller noe om hvorvidt det er konflikt i et land det inneværende året, men fanger opp utbrudd av nye konflikter. Dette kan man for eksempel se ved å undersøke Rwanda, som i datasettet er kodet med «1» i 1990 da konflikten startet, og deretter i årene frem til 1994, når konflikten ender, er kodet med «0». På den måten kan man sjekke ut om befolkningsrike land har flere utbrudd av konflikt uavhengig av hvor lenge konfliktene varer. Varighet på konflikt er i seg selv interessant, og det kan stilles spørsmål ved om befolkningsrike land også opplever lengre konflikter internt, men det er ikke fokus for denne oppgaven.

Valget om å bruke UCDP sin definisjon på intern konflikt er tatt med tanke på problemstillingen. Grunnen til dette er at andre datasett for eksempel opererer med et høyere dødstall før det regnes som konflikt, slik som blant annet Correlates of war (COW) gjør, der det må være minst 1000 døde i et kalenderår før det regnes som konflikt. Ved å bruke Uppsala sin definisjon, står man ikke i fare for at konflikter i små land blir ekskludert, slik man kan risikere med COW sin definisjon. For denne analysen er det helt avgjørende at dataene og definisjonene som brukes ikke favoriserer konflikter i land med stor befolkning, da dette kan farge resultatene i favør av befolkningsrike land allerede før analysen er gjort. Ved å bruke en definisjon som inkluderer konflikter med 25 kamprelaterte dødsfall vil man med stor sannsynlighet fange opp konflikter av relevans i alle land, uavhengig av befolkningsstørrelse.

Når det gjelder å definere hva intern konflikt egentlig er, finnes det ulike definisjoner. En slik forskjell ser man blant annet mellom datasettet til UCDP/PRIO og COW. Som jeg allerede har påpekt i denne delen av oppgaven, vil bruken av de ulike datasettene potensielt kunne gi ulike svar. Der det ene datasettet vil definere en situasjon som en intern konflikt, vil situasjonen i andre datasett ikke nå opp til definisjonen som brukes. Scarcelli (2017) eksemplifiserer dette ved å belyse situasjonen i Irak og hvorvidt konflikten er intern, mellomstatlig eller begge deler. Her kommer Scarcelli (2017:98) frem til at det er datasettet til UCDP/PRIO som er nærmest å løse dette problemet. Krigen i Irak defineres av UCDP/PRIO som mellomstatlig i 2003, fulgt av en internasjonalisert intern konflikt fra 2004 frem til USA

forlater Irak i 2011, før konflikten defineres som intern i to år frem til USA igjen inkluderes i det IS invaderer landet i 2014 (Scarcelli, 2017:98). Andre datasett vil definere denne perioden og disse hendelsene annerledes.

3.2.1 Temporal avhengighet

Observasjonene som utgjør datagrunnlaget for den avhengige variabelen, konflikt, er temporal avhengige. Det vil si at verdier for ett år i et land, er avhengig av verdier fra tidligere år. Dette er også testet empirisk, der man kan se at konflikthistorie påvirker sannsynlighet for ny konflikt (Beck, Katz & Tucker, 1998:1260). Temporal avhengighet bryter forutsetningen om at observasjonene skal være uavhengige av hverandre, og kan slå ut som overoptimistisk inferens, der variasjonen blir underestimert som fører til at t-verdiene blir større enn de egentlig skal være (Bech, Katz & Tucker, 1998:1261). Siden variabelen er dikotom, må den temporale avhengigheten kontrolleres for på andre måter enn å «lagge» den. Dette gjøres ved å inkludere en variabel som kontrollerer for antall år siden siste konflikt ved hjelp av «btsacs», programmert av Tucker (1999). Variabelen «fredsår» med de tre tilhørende _spline-variablene kontrollerer tidsavhengighet og gjør t-verdiene mer korrekte.

For å kunne konstruere «fredsår»-variabelen måtte det tas utgangspunkt i UCDP/PRIO Armed Conflict 19.1, og ikke UCDP/PRIO Onset 19.1, ettersom dette datasettet ikke inkluderer noe informasjon om hvor lenge det er siden forrige konflikt. I Armed Conflict 19.1 er hvert år med konflikt kodet som konflikt uavhengig av om det gjelder utbrudd eller videreføring av eksisterende konflikt. Da disse to datasettene ble satt sammen, viste det seg at Onset-datasettet hadde 32 observasjoner av intern konflikt som ikke var kodet som dette i Armed Conflict-datasettet. Dette er i grunn rart, ettersom Onset-datasettet tar utgangspunkt i Armed Conflict-datasettets type 3 og 4 konflikter. Her har jeg derfor valgt å kode om de 32 observasjonene som kommer fra Onset-datasettet fra «1» til «0», slik at dataene som brukes i hovedsak følger Armed Conflict-datasettet. De 32 observasjonene dette gjelder inkluderes i appendiks.

3.3 Uavhengige variabler (X)

For å kunne undersøke hvorfor befolkningsrike land oftere opplever utbrudd av intern konflikt enn land med mindre befolkning, inkluderes det flere ulike uavhengige variabler som kan tenkes å forklare den signifikante og robuste påvirkningen. Disse variablene måler ulike aspekter ved land, og kan på ulike måter relateres til befolkningsstørrelse. I følgende delkapittel blir disse operasjonaliseringene gjennomgått.

3.3.1 Befolkningsstørrelse

Befolkningsstørrelse, som er hovedvariabelen som testes, er hentet fra Verdensbanken, og måler befolkning utfra definisjonen på populasjon, som inkluderer alle borgere i landet uavhengig av lovlig status eller statsborgerskap. Verdiene for hvert år er målt midt i året, og er vanligvis basert på nasjonale folketellinger. I utviklingsland kan tallene være mer eller mindre korrekte på grunn av begrensninger med tanke på transport og kommunikasjon som trengs for å samle inn dataene. Kvaliteten og reliabiliteten til offisielle demografiske data er også påvirket av samfunnets tillit til regjeringen, konfidensialitet og beskyttelse mot misbruk av data. Populasjonsindikatoren kan også være begrenset av ulike konseptualiseringer, definisjoner og innsamlingsmetoder i hvert enkelt land. Internasjonal migrasjon er den eneste andre faktoren i tillegg til fødsels- og dødsrater som direkte påvirker et lands befolkningsvekst, og denne faktoren vil det være vanskelig å estimere. Standarder for varighet og intensjon for internasjonale flyttinger som kvalifiserer som migrasjon varierer, og man kan forvente noe ulike tall basert på dette (World bank, 2020). Dataene som Verdensbanken står ansvarlig for er samlet inn gjennom et samarbeid mellom 6 ulike aktører, og har id SP.POP.TOTL. Variabelen har også blitt log-transformert for å rette opp skjevheten i observasjonene.

3.3.2 Landareal

Dataene for hvert enkelt lands areal er hentet ut av Verdensbankens datasett (World bank, 2020). Variabelen måler et lands totale areal i kvadratkilometer, ekskludert vannområder, nasjonale krav på kontinentalsokkelen og eksklusive økonomiske soner. I de fleste tilfeller inkluderer definisjonen av vannområder store elver og innsjøer innad i landene. Verdensbanken har hentet dataene fra «Food and Agriculture Organization» (FAO), og har id AG.LND.TOTL.K2. FAO prøver å implementere en standard definisjon og standardmetoder for rapportering, men det vil likevel kunne være små ulikheter over tid og mellom land. Dataene som er samlet fra offisielle nasjonale kilder er supplementert med offisielle data fra sekundærkilder. Sekundærkildene inkluderer offisielle landdata fra internettsider, nasjonale publikasjoner og relaterte landdata rapportert fra ulike internasjonale organisasjoner (World bank, 2020). Variabelen har også blitt log-transformert for å rette opp skjevheten i observasjonene.

3.3.3 Rural befolkning

Dataene som måler rural befolkning refererer til mennesker som lever i rurale områder som definert av nasjonale statistikere midt i det aktuelle året, og er hentet fra Verdensbankens datasett med id SP.RUR.TOTL (World Bank, 2020). Dataene er kalkulert som forskjellen mellom andel av befolkningen som bor i urbane områder og den totale befolkningen. Det er ingen universell standard for hvordan man skiller rurale fra urbane områder, og alle urban-rural dikotomier er en oversimplifisering av dette forholdet. Ofte brukes en operasjonell definisjon av ruralitet basert på befolkningstetthet og distanse til store byer (Chomitz, Buys & Thomas, 2005). Mange land bruker også en urban klassifisering relatert til størrelsen eller karakteristikken på bosettingsmønsteret i området. Noen definerer urbane områder basert på om de har spesifikk infrastruktur eller tilbud til befolkningen. På grunn av nasjonale ulikheter i definisjonen på urbane områder, er det vanskelig å komme med en enkelt definisjon på denne variabelen. For noen land vil det for eksempel ikke gi mening å slå sammen urban og rural befolkning, ettersom dette ikke gir det totale befolkningstallet i landet. Observasjonene på denne variabelen vil altså ikke være presise, men vil kunne gi et inntrykk av hvordan befolkningen fordeler seg utover landets totale areal, og da særlig i de mindre utbyggede områdene (World Bank, 2020). Variabelen er konstruert ved å dividere den opprinnelige variabelen på hundre og multiplisere den med total befolkning, slik at variabelen har gått fra å måles i prosent, til å måle totalt antall. Denne variabelen er også log-transformert.

3.3.4 Ungdomsbølge

Denne variabelen viser til totalt antall unge menn mellom 25 og 29 år i et land. Variabelen baserer seg på total mannlig befolkningsstørrelse og prosentvis mannlig befolkning mellom 25 og 29 år i landet. Variabelen har blitt konstruert ved å dividere den opprinnelige variabelen for unge menn med hundre, og multiplisere dette med den totale mannlige befolkningsstørrelsen, slik at variabelen måles i totalt antall unge menn. Estimaten er basert på fordelingen av kjønn og alder fra United Nations Population Division's World Population Prospects, og er hentet fra Verdensbanken med id SP.POP.2529.MA.5Y (World Bank, 2020). Variabelen er også log-transformert.

3.3.5 Totalt bruttonasjonalt produkt

Det totale bruttonasjonale produktet for et land er summen av alle bruttoverdier for alle bosatte produsenter i økonomien, pluss eventuelle produktskatter og minus eventuelle subsidier som ikke er inkludert i verdien av produktene. Det beregnes uten å gjøre fradrag for

avskrivning av fabrikkerte eiendeler eller for uttømming og forringelse av naturressurser. Dataene er hentet fra Verdensbanken, med id NY.GDP.MKTP.CD (World Bank, 2020). Dataene er oppgitt i nåværende amerikanske dollar. Målene i dollar for BNP blir konvertert fra innenlandske valutaer ved å bruke offisielle valutakurser for ett år. For noen få land der den offisielle valutakursen ikke gjenspeiler kursen som brukes på faktiske valutatransaksjoner, brukes en alternativ konverteringsfaktor. Ulike land bruker ulike definisjoner, metoder og rapporteringsstandarder når de viser til sin totale BNP. Verdensbanken gjør noen ganger endringer i dataene som er rapportert inn fra hvert enkelt land for å forbedre dataene etter internasjonale retningslinjer. Dette er gjort fordi mange land, spesielt utviklingsland, har store begrensninger i ressurser, tid, trening og budsjetter som kreves for å produsere troverdige nasjonale statistikker (World bank, 2020). Variabelen har også blitt log-transformert for å rette opp skjevheten i observasjonene.

3.3.6 Etnopolitiske relevante grupper

Denne variabelen viser til antall etnopolitiske relevante grupper i hvert land, og er hentet fra Ethnic power relations-datasettet (EPR) (Vogt m.fl., 2015). En etnisk gruppe er definert som en gruppe som subjektivt føler seg samlet rundt et felles opphav eller delt kultur (Vogt m.fl., 2015). Ulike markører kan indikere et slikt delt opphav eller kultur. Definisjonen som brukes her er at etnisitet inkluderer språk, hudfarge og religion. En etnisk relevant gruppe er en gruppe hvor minst en signifikant politisk aktør hevder å representere interessene til en etnisk gruppe på nasjonalpolitisk arena. Med en signifikant politisk aktør menes et medlem av en politisk organisasjon som er aktiv i den nasjonale politikken (Vogt m.fl., 2015).

3.3.7 Ekskluderte etnopolitiske grupper

Variabelen viser til antall politisk ekskluderte etniske grupper i det respektive landet, og er hentet fra EPR-datasettet (Vogt m.fl., 2015). Etniske grupper defineres på samme måte som i 3.3.6. Politisk eksklusjon må være direkte rettet mot en etnisk gruppe, og ser bort fra indirekte diskriminering som færre muligheter til utdanning eller mindre tilgang til arbeidsmarkedet (Vogt m.fl., 2015). Ekskluderingen må være aktiv, intensjonell og med et mål om å ekskludere gruppen fullstendig fra den politiske sfæren.

3.3.8 Politisk korrupsjon

Spørsmålet variabelen måler er «hvor overgripende er politisk korrupsjon?», og er hentet fra Varieties of democracy-datasettet (V-dem) (Coppedge m.fl., 2019a). Den er målt på

intervallnivå fra lav til høy (0-1) (Ringdal, 2013:91). Indeksen inkluderer målinger på seks ulike typer korrupsjon som omhandler både ulike arenaer og nivåer av politikken, og skiller mellom lovgivende, dømmende og utøvende korrupsjon. Innenfor utøvende korrupsjon skiller det mellom korrupsjon med mål knyttet til bestikkelse eller underslag. Den skiller også på om korrupsjonen er på høyere eller lavere nivå innad i den utøvende makten. Indeksen består av gjennomsnittet av (a) offentlig sektor korrupsjonsindeks, (b) utøvende makt korrupsjonsindeks, (c) indikatoren for lovgivende maktkorrupsjon, og (d) indikatoren for dømmende maktkorrupsjon. Med andre ord er de fire ulike målingene på korrupsjon i ulike deler av staten vektet likt (Coppedge m.fl, 2019b). Variabelen manglet et par observasjoner, så den har blitt inter- og ekstrapolert.

3.4 Kontrollvariabler

Det kontrolleres for to variabler i modellene som blir kjørt for å teste hypotesene. Den første variabelen er «fredsår» som kontrollerer for hvor mange år det er siden forrige konflikt. Den andre variabelen det kontrolleres for er BNP per capita, ettersom dette er en ofte brukt kontrollvariabel sammen med befolkning.

3.4.1 Fredsår

Variabelen «fredsår» er konstruert med utgangspunkt i UCDP/PRIO Armed Conflict 19.1, ved hjelp av «btses», som forklart under delkapittelet om temporal avhengighet. Det er konflikter kodet med «3» og «4» i datasettet som har vært utgangspunkt for konstruksjonen. Disse konfliktene er definert som «(3) interne konflikter der A alltid er en stat, mens B enten alltid er minst en oppgjørsggruppe eller (4) at det er minst en opprørsgruppe i tillegg til innblanding fra internasjonale tropper» (Pettersen, 2019). Variabelen måler antall år siden siste konfliktutbrudd.

3.4.2 Bruttonasjonalt produkt per capita

Bruttonasjonalt produkt per capita (BNP per capita) er hentet fra Verdensbanken med id NY.GDP.PCAP.CD, og er definert som det bruttonasjonale produktet delt på total befolkning i midten av året (World Bank, 2020). Det totale bruttonasjonale produktet for et land er summen av alle bruttoverdier for alle bosatte produsenter i økonomien pluss eventuelle produktskatter og minus eventuelle subsidier som ikke er inkludert i verdien av produktene. Det beregnes uten å gjøre fradrag for avskrivning av fabrikkerte eiendeler eller for uttømming

og forringelse av naturressurser. Data er i nåværende amerikanske dollar. Variabelen har også blitt log-transformert for å rette opp skjevheten i observasjonene.

3.5 Deskriptiv statistikk

	N	gjennomsnitt	standardavvik	min	max
År	7123	1992.892	15.64937	1961	2017
Konfliktutbrudd	7123	.0360803	.1865031	0	1
(ln)Befolkning	7123	15.98605	1.610695	12.09531	21.04997
(ln) Unge menn	7123	7.811738	1.616786	8.656432	17.99915
(ln)Rural befolkning	7123	49.17818	2.289959	0	20.60546
Politisk korrupsjon	7123	.4885757	.3076129	.006	.98
Ekskluderte grupper	7123	2.230802	4.180204	0	55
Politisk relevante grupper	7123	3.971641	4.722818	0	57
(ln)BNP	7123	23.47239	2.304746	17.41914	30.60069
(ln)BNP per capita	7123	8.18778	1.517026	4.892626	11.66336
(ln) Landareal	7123	12.13104	2.006449	5.70711	16.61218
Fredsår	7123	17.42524	16.03825	0	57
_spline1	7123	-2398.6	3194.804	-12320	0
_spline2	7123	-8997.901	13228.68	-52118	0
_spline3	7123	-10224.36	16983.07	-70928	0

Tabell 3: Deskriptiv statistikk for variablene

3.6 Metode

I denne delen av oppgaven vil fordelene og ulempene ved bruk av paneldata, som er designet på datamaterialet som brukes i den kommende analysen, bli gjennomgått. Deretter vil det bli redegjort for hvordan manglende verdier har blitt håndtert, før det følger en kort diskusjon rundt dataenes reliabilitet og validitet med utgangspunkt i ulike innsamlingsmetoder for ulike datasett og operasjonaliseringer av de ulike teoretiske argumentene som danner grunnlag for analysen. Etersom oppgaven bruker logistisk regresjon som metode, følger også en kort gjennomgang av fordelene og ulempene ved denne analysemetoden, før det avslutningsvis blir gått igjennom relevante diagnosetester som er gjort. Det har blitt testet for multikollinearitet i modell 1-8, korrelasjon mellom variablene, og det har blitt sjekket for ekstremverdier og uteliggere. Her har det vært modell 1, med befolkningsstørrelse og fredsår, som har blitt testet.

3.6.1 Paneldata

I denne studien finnes det 7123 observasjonstidspunkter fra 1961 til 2017. Designet for studien er et prospektivt longitudinelt design, som vil si at observasjonene har startet på et gitt tidspunkt og deretter har observasjonsenhetene blitt fulgt fremover i tid. I longitudinelle design følger man de samme observasjonsenhetene over tid. Longitudinelle studier kalles også panelstudier (Skog, 1998:70). Utvelgelsen av data er basert på hvilke land og årstall som har observasjoner på de ulike variablene som brukes som uavhengige variabler. Det er ikke trukket ut et tilfeldig utvalg, ettersom studien er spesielt opptatt av virkningene av bestemte typer påvirkning som observasjonsenhetene utsettes for, og utvelgingene har blitt foretatt med dette som utgangspunkt.

Paneldata er altså en multidimensjonal type data, i motsetning til tverrsnitt eller tidsserier som bare har en dimensjon (Mehmetoglu & Jakobsen, 2017:228). Det vil si at dataene både kan si noe om forskjeller mellom de ulike landenes målinger på enkeltvariabler, men også hvordan disse forskjellene eventuelt endrer seg over tid. I følge Baltagi (2008), er det også syv andre fordeler med å bruke denne type data: (1) Paneldata kontrollerer for individuell heterogenitet. Man anerkjenner dermed at individuelle enheter kan ha tidskonstante egenskaper ved seg som kan ha en effekt på den avhengige variabelen. (2) Paneldata gir mer informasjon, mer variabilitet, mindre kollinearitet mellom variablene, flere frihetsgrader og mer effektivitet. (3) Paneldata fanger bedre opp dynamikken i samfunnet. Eksempel på dette kan være fattigdom. (4) Paneldata fanger opp og identifiserer effekter som ikke kan oppdages i vanlig tverrsnitt eller tidsserier. (5) Makro paneldata har en lengre tidsserie (Baltagi, 2008:4-7).

Selv om det er mange fordeler, finnes det også et par ulemper med bruk av paneldata. Som allerede gjort rede for i gjennomgangen av variablene, kan det være vanskelig å sikre god datainnsamling over lengre tid, og spesielt på tvers av land. Det er kanskje denne ulempen som har størst påvirkning på denne analysen, ettersom det er mange land og mange variabler spredt utover flere år som danner grunnlaget for studien. Likevel er dataene som har blitt brukt i denne studien hentet fra anerkjente aktører, som også har diskutert og vurdert svakhetene ved egen datainnsamling.

Paneldata vil også ha utfordringer med korrelerte tidsserier og temporal avhengighet, som tidligere er diskutert med tanke på den avhengige variabelen i analysen. Når man måler enheter over lengre tid, vil observasjonene ofte være påvirket av observasjonen på det forrige

tidspunktet den ble målt. Antall unge menn i Afghanistan vil for eksempel ikke endre seg drastisk fra ett år til et annet, og det samme gjelder for majoriteten av variablene som er målt, og brukes, i denne analysen. Det gjør at man kan forvente autokorrelasjon i residualene. Når seriene først er på et høyt eller lavt nivå, har de altså en tendens til å forbli på dette nivået en stund, et fenomen som kalles «random walk»-serie (Skog, 2017:331). En høy autokorrelasjon mellom observasjonene vil kunne påvirke standardfeilen til koeffisientene i analysene, noe det er viktig å være bevisst på (Skog, 2017:333).

3.6.2 Håndtering av manglende verdier

For å håndtere manglende verdier i datagrunnlaget er det spesielt tre tiltak som har blitt gjort. Det første er å interpolere og ekstrapolere variabler med noen manglende verdier, som her først og fremst gjelder korrupsjonsvariabelen. Interpoleringen predikerer verdier mellom faktiske observasjoner, mens ekstrapoleringen gjør det mulig å predikere verdier også uten å ha faktiske observasjoner på hver side.

For det andre kjøres alle regresjonsanalysene på det samme datasettet. Dette gjøres for å sikre at alle regresjonene har det samme grunnlaget. Om det ikke hadde blitt gjort kunne noen av analysene som har blitt kjørt vært på mange flere eller færre observasjoner enn hva de andre analysene er basert på, noe som ville gjort det vanskelig å sammenlikne de ulike effektene. For å unngå dette har det blitt kjørt en regresjonsanalyse som inneholder alle variablene som er aktuelle for oppgaven. Deretter har datagrunnlaget blitt endret gjennom kommandoen «keep if e(sample)», slik at datagrunnlaget endres til å kun inneholde observasjoner som har verdier på alle variablene som er inkludert i modellene.

I utvelgelsen av relevante variabler som best kunne operasjonalisere det teoretiske grunnlaget, har det også blitt tatt hensyn til hvorvidt det er mye manglende data. Det er spesielt to variabler som har blitt valgt bort på bakgrunn av dette, som ideelt sett kunne vært gode indikatorer. Som nevnt tidligere i teorigjennomgangen, er det vanlig å måle etnisk mangfold med en ELF-indeks. Denne indeksen har kun data frem til 2005, noe som er utfordrende med tanke på at analysene kjøres på data frem til 2017. I tillegg mangler indeksen observasjoner på ulike årstall for de ulike landene. Interpolering og ekstrapolering har blitt brukt for å fylle inn de manglende verdiene på variabelen, men etter en gjennomgang av observasjonene kan man se at resultatet av dette ikke gir optimale verdier, samtidig som det fortsatt er mange land som mangler observasjoner for alle årstallene, og dermed ikke kan være med i analysen. I stedet

har antall relevante etniske grupper og antall ekskluderte grupper blitt brukt for å måle etnisk mangfold. Det samme gjelder variablene som var ment å måle andel fattige i et land. Her har både GINI-koeffisienten, relativ fattigdom og absolutt fattigdom blitt forsøkt å bruke som variabler. I disse variablene har det også vært problemer med mye manglende verdier som gjør det vanskelig å bruke. Noen land har ingen målinger på noen av verdiene, mens andre har målinger på tilfeldige år. Derfor har også disse variablene blitt vurdert til å ikke ha et godt nok datagrunnlag til å inkluderes i analysen. BNP per capita har vært det eneste målet med et tilfredsstillende antall observasjoner, men her igjen er det et dårlig mål på antall fattige i et land, da det bare sier noe om gjennomsnittet, og ikke hvordan pengene er distribuert blant befolkningen. Et mål på generell fattigdom har derfor blitt utelatt, og det har heller blitt lagt fokus på utsatte grupper i samfunnet som også generelt kan sies å være blant de fattigste.

3.6.3 Validitet og reliabilitet

Reliabilitet, eller pålitelighet, går på om gjentatte målinger med samme måleinstrument gir samme resultat. Validitet, eller gyldighet, går på om en faktisk måler det en vil måle. Denne typen validitet kalles også begrepsvaliditet. Validitet er det mest generelle av de to begrepene. Høy reliabilitet er en forutsetning for høy validitet (Ringdal, 2013:96). Man kan si at reliabilitet er et rent empirisk spørsmål, mens validitet krever en teoretisk vurdering, ettersom det alltid må referere til den teoretiske sammenhengen begrepet brukes i (Ringdal, 2013:97).

Reliabilitet påvirkes av tilfeldige målefeil, mens systematiske målefeil går direkte ut over dataenes validitet. Allmenn kildekritikk er en måte å vurdere dataenes reliabilitet på, der en gjennomgang av hvordan spørsmålene er formulert, registrert og samlet inn gjør det mulig å spore mulige feilkilder (Ringdal, 2013:97). Når det gjelder dette datasettets reliabilitet er alle variablene hentet fra anerkjente kilder som har veldokumenterte innsamlingsmetoder og redegjørelser for eventuelle svakheter ved datasettene. Både UCDP/PRIO, WBI, EPR og V-dem regnes som troverdige kilder, men fordi innsamlingen av data for de ulike variablene er gjort av en rekke forskjellige mennesker i forskjellige land, kan dette påvirke datagrunnlagets reliabilitet. I noen land kan de systematisk ha blitt oppjustert, mens de i andre land er nedjustert. På denne måten kan man aldri være helt sikker på at vi målingene stemmer med virkeligheten. Total BNP kan fungere som et eksempel her. Om noen land har som hensikt å virke sterkere enn de egentlig er, kan de ha justert opp sin totale BNP. Når det gjelder populasjon, kan folketallet variere også utfra hvor enkelt det er å telle antall mennesker i landet. Særlig utviklingsland kan ha problemer med å komme til i alle områdene av landet, og

målingene kan derfor bli noe usikre. Verdensbanken, som har ansvar for disse målingene har hatt muligheten til å justere på verdiene til de ulike variablene, som gjør at dataenes validitet styrkes.

Konfliktvariabelens reliabilitet bør være høy, da både PRIO og UCDP er ledende forskningsinstitutter når det kommer til innsamling av data og utvikling av feltet. De har argumentert for reliabiliteten til dataene deres på en overbevisende måte, og innsamlingsprosedyrene er veldokumentert. For årene før 1989, er dataene sjekket gjennom en tre-steps prosess (Gleditsch m.fl., 2002:618). Det har blitt samlet informasjon fra et flertall av eksisterende datasett, før de kritisk har valgt ut hvilke potensielle konflikter som har passet inn i deres definisjoner av konflikt. Deretter har alle land blitt undersøkt for å avdekke eventuelle konflikter som ikke er medberegnet i allerede eksisterende datasett ved å gjøre dybdestudier i land der det muligens har vært konflikter. Ifølge Gleditsch med flere (2002), er det liten sannsynlighet for at det er blitt kodet konflikt i land der det ikke har vært konflikt, mens det er en litt større sannsynlighet for at det kan være flere konflikter i land som de ikke har fått med seg. Om man tar utfordringene med å kode historiske data i betraktning, kan man likevel vurdere reliabiliteten til datasettet til UCDP/PRIO som tilfredsstillende.

Begrepsvaliditet handler om hvorvidt man faktisk måler det teoretiske begrepet vi ønsker å måle (Ringdal, 2013:98). Det er altså relasjonen mellom indikatorene og det teoretiske begrepet som undersøkes. Det vil sjeldent være full validitet i en oppgave som denne, ettersom det vil være umulig å fange opp alle aspektene ved de teoretiske begrepene som brukes. Det vil sjeldent være fullt samsvar mellom teoretisk og operasjonell validitet, så spørsmålet er heller om indikatorene som brukes gir en rimelig dekning av en hypotetisk populasjon av indikatorer, det vil si om målet dekker de viktigste aspektene av begrepet (Ringdal, 2013:98). En slik avgjørelse er på mange måter subjektiv, og det er opp til forskeren selv å avgjøre om variablene måler det de er valgt ut til å måle. For at studien skal ha en høyere validitet har jeg, der den teoretiske gjennomgangen viser behov for det, inkludert flere ulike variabler som skal måle noe av det samme før det har blitt tatt en vurdering på hvilken variabel som best måler det aktuelle aspektet. Når det kommer til korrupsjon, brukes det for eksempel en indeks bestående av fire ulike aspekter ved korrupsjon til å operasjonalisere det teoretiske argumentet, og sannsynligheten for at de viktigste aspektene da er dekket, vil være økt.

3.6.4 Logistisk regresjon

Denne oppgaven bruker logistisk regresjon for å teste hypotesene som er konstruert om hvorfor befolkningsstørrelse gir større sannsynlighet for utbrudd av konflikt. Siden konflikt er en dikotom variabel der verdiene enten er 0 eller 1, altså «ikke utbrudd» eller «utbrudd», egner logistisk regresjon seg som analyseverktøy. Dette gjøres fordi et av kravene til lineær regresjon er at den avhengige variabelen er kontinuerlig (Mehmetoglu & Jakobsen, 2017:161). I denne oppgavens analyser er ikke dette tilfellet, ettersom vi ønsker å undersøke om et fenomen (konflikt) er til stede eller ikke. Ved å bruke lineær regresjon til å undersøke en dikotom variabel, vil det være flere problemer. For det første vil det ikke være mulig å lage en lineær linje mellom X og Y, og for det andre vil det være heteroskedastisitet til stede i modellene (Mehmetoglu & Jakobsen, 2017:161). For det tredje risikerer man å predikere verdier som faller utenfor intervallet mellom 0 og 1, når det i realiteten ikke finnes andre verdier enn disse i den avhengige variabelen.

Ved å bruke logistisk regresjon i stedet for lineær regresjon, bortfaller en del av disse problemene. I lineær regresjon estimeres det hvor mye den avhengige variabelen endres når den uavhengige variabelen endres med 1. I logistisk regresjon estimerer man hvor mye den naturlige logaritmen for oddsene for at $Y=1$ endrer seg for hver endring i X. Odds er en annen måte å uttrykke sannsynlighet på (Mehmetoglu & Jakobsen, 2017:163). I stedet for at modellen søker etter en lineær linje, vil det ved logistisk regresjon danne seg en s-formet kurve som flater ut når den nærmer seg 0 eller 1, ettersom verdier over dette ikke vil være mulig. I regresjonsmodeller der Y har verdiene 0 og 1 kan de predikerte y-verdiene tolkes som sannsynligheter og regresjonskoeffisientene tolkes som effekter av x-variablene på sannsynligheten for at Y skal være 1 (Ringdal, 2013:435).

I denne oppgaven er det regresjonskoeffisientene som vil bli brukt til å tolke de uavhengige variablenes effekt på sannsynligheten for at det skjer et utbrudd av konflikt. Hovedoppgaven vil være å finne ut hva som kan forklare effekten som befolkningsstørrelse har på denne sannsynligheten. Ved å inkludere en enkelt variabel i regresjonsanalysen sammen med befolkningsstørrelse, vil vi se etter variabler som minker sannsynligheten for at stor befolkning gir konfliktutbrudd. De variablene som minker denne sannsynligheten når de inkluderes, vil være variabler som kan være med på å forklare hvorfor befolkningsstørrelse har en slik påvirkning på konflikt.

3.6.5 Diagnosetesting

For at resultatene av denne analysen skal være mest mulig troverdig, er det viktig å sette modellen i et kritisk lys, slik at eventuelle svakheter blir synlig. Ofte kan svakheter rettes opp i, mens det andre ganger vil være viktigst å få frem hvilke begrensninger de eventuelle svakheter legger på modellen og dens forklaringssevne (Midtbø, 2016). I denne delen av metodekapittelet vil det bli gjennomgått og sjekket for korrelasjon mellom variabler og innad i modeller, ekstremverdier og uteliggere. Multikollinearitet blir sjekket fordi det er viktig for oppgavens problemstilling at de ulike regresjonskoeffisientene gir mening. Korrelasjonen mellom befolkningsstørrelse og de andre uavhengige variablene er viktig for oppgaven fordi det uten en korrelasjon med befolkning vil gi lite mening å anta at de kan forklare sammenhengen mellom befolkning og konflikt. Til sist vil ekstremverdier og uteliggere sjekkes for fordi usedvanlige observasjoner som har mye å si for resultatene i analysene kan gi oss et feil inntrykk av koeffisientene, og potensielt gi oss resultater som ikke gjenspeiler de egentlige tendensene i datasettet. Koeffisienter som står og faller på en eller noen få verdier har begrenset troverdighet, ettersom den estimerte effekten da forutsetter tilstedeværelsen av en bestemt enhet, og ikke har noen generell gyldighet (Midtbø, 2016:115).

3.6.5.1 Multikollinearitet

En regresjonsanalyse forutsetter at forklaringsvariablene ikke skal være perfekt kollineære, som de er hvis en variabel er en nøyaktig lineær kombinasjon av de andre. Ved perfekt multikollinearitet er det umulig å identifisere regresjonskoeffisientene. Det er bare perfekt, ikke høy, multikollinearitet som bryter med regresjonsforutsetningene (Midtbø, 2016:128). Når det er sagt, vil høy multikollinearitet gjøre det vanskelig å beregne størrelsen på de ulike koeffisientene, og ofte vil små endringer i modellen føre til store endringer i resultatene. Modellene er dessuten ofte kjennetegnet ved en høy R^2 og høye standardfeil (Midtbø, 2016:128). Dette er verdt å merke seg ettersom denne oppgaven søker å nettopp se etter endring i koeffisienten til befolkningsstørrelse når man inkluderer andre variabler. Hvis det er høy multikollinearitet mellom de ulike variablene vil det også gjøre det vanskelig å måle hvilke variabler som har noe å si, og hvilke som ikke har det. I denne oppgaven er det ønskelig med høy korrelasjon mellom befolkningsstørrelse og de aktuelle variablene, og lav korrelasjon mellom de andre variablene. Derfor forventes det høy multikollinearitet i noen av modellene.

I appendiks ligger tabellen som viser VIF mellom alle variablene som er inkludert i de ulike modellene. VIF er definert som $1/(1-R^2)$ i en regresjon der en forklaringsvariabel blir forklart av de andre forklaringsvariablene (Midtbø, 2016:129). Et ekvivalent mål er toleransen definert som $1/VIF$. Det er vanskelig å fastslå hva som er for mye multikollinearitet, men ifølge en tommelfingerregel er VIF større enn 10, eller en toleranse på under 0,10 problematisk (Midtbø, 2016:129). Det hevdes også av flere at problemet med multikollinearitet er overdrevet, og at det ikke utgjør et så stort problem som forespeilet (Midtbø, 2016:129). Det er spesielt modell 5 og 8 som lider av høy multikollinearitet, der variablene «unge menn» og «BNP» er inkludert. I modell 8 er det høy korrelasjon mellom befolkning og unge menn med en VIF på 82, mens de andre inkluderte variablene ikke korrelerer særlig med hverandre. I modell 8 har «BNP» en VIF på 14, og både «befolkning» og «bnp per capita» har en litt forhøyet VIF, selv om den ikke er over 10.

3.6.5.2 Korrelasjonsanalyse

For å kunne svare på problemstillingen om hva som gjør at befolkningsrike land opplever mer borgerkrig enn andre, har det vært nødvendig å finne mulige forklaringsfaktorer som både korrelerer med befolkningsstørrelse og med konflikt. Samtidig er det ønskelig at variablene som skal forklare sammenhengen mellom konflikt og befolkningsstørrelse ikke korrelerer for mye med hverandre. Ved å bruke Pearsons R måler man den lineære sammenhengen mellom to kontinuerlige variabler, og det er dette som har blitt gjort for å sjekke sammenhengen mellom befolkningsstørrelse og de mulige forklaringsvariablene. Når Pearsons R er mindre enn null går lave verdier på den ene variabelen sammen med høye verdier på den andre variabelen, og motsatt (Midtbø, 2016:84). Jo nærmere R ligger -1 eller +1, desto sterkere er korrelasjonen (Midtbø, 2016:84).

I teorikapittelet har korrelasjonen mellom de ulike variablene og befolkningsstørrelse allerede blitt vist, mens man her ser sammenhengen i tall. Som en kan lese ut av korrelasjonsmatrisen i Tabell 4, er det en del variabler som korrelerer sterkt med befolkning. Disse beskriver deler av den totale befolkningen, så denne korrelasjonen er naturlig sterk. Dette kan man for eksempel se med variablene som måler andel unge menn av total mannlig populasjon og størrelse på den rurale befolkningen i landet. I følge Mehmetoglu og Jakobsen er en korrelasjon på 0,1, 0,3 og 0,5 henholdsvis svak, moderat og sterk (2017:39). Det er en svak korrelasjon mellom befolkning og korrupsjon, men denne inkluderes likevel fordi korrupsjon er med på å støtte

opp argumentet om at høy BNP gjør det mer fristende å ta over et land. Resten av variablene har enten en moderat eller sterk korrelasjon med befolkningsstørrelse.

	Konflikt	(Ln) befolkning	(Ln) areal	Relevante grupper	Ekskluderte grupper	(Ln) rural befolkning	(Ln) unge menn	(Ln) BNP p.cap	Korrupsjon	(Ln) BNP
Konflikt	1									
(Ln) befolkning	0.131***	1								
(Ln) areal	0.102***	0.677***	1							
Relevante grupper	0.144***	0.454***	0.389***	1						
Ekskluderte grupper	0.094***	0.418***	0.356***	0.903***	1					
(Ln) Rural befolkning	0.122***	0.746***	0.632***	0.364***	0.344***	1				
(Ln) unge menn	0.130***	0.994***	0.665***	0.455***	0.419***	0.725***	1			
(Ln) bnp p.cap	-0.098***	-0.046***	-0.089***	-0.182***	-0.093***	-0.321***	-0.021	1		
Korrupsjon	0.088***	0.117***	0.152***	0.140***	0.101***	0.284***	0.117***	-0.646***	1	
(Ln) BNP	0.027*	0.685***	0.381***	0.215***	0.239***	0.306***	0.704***	0.636***	-0.331***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

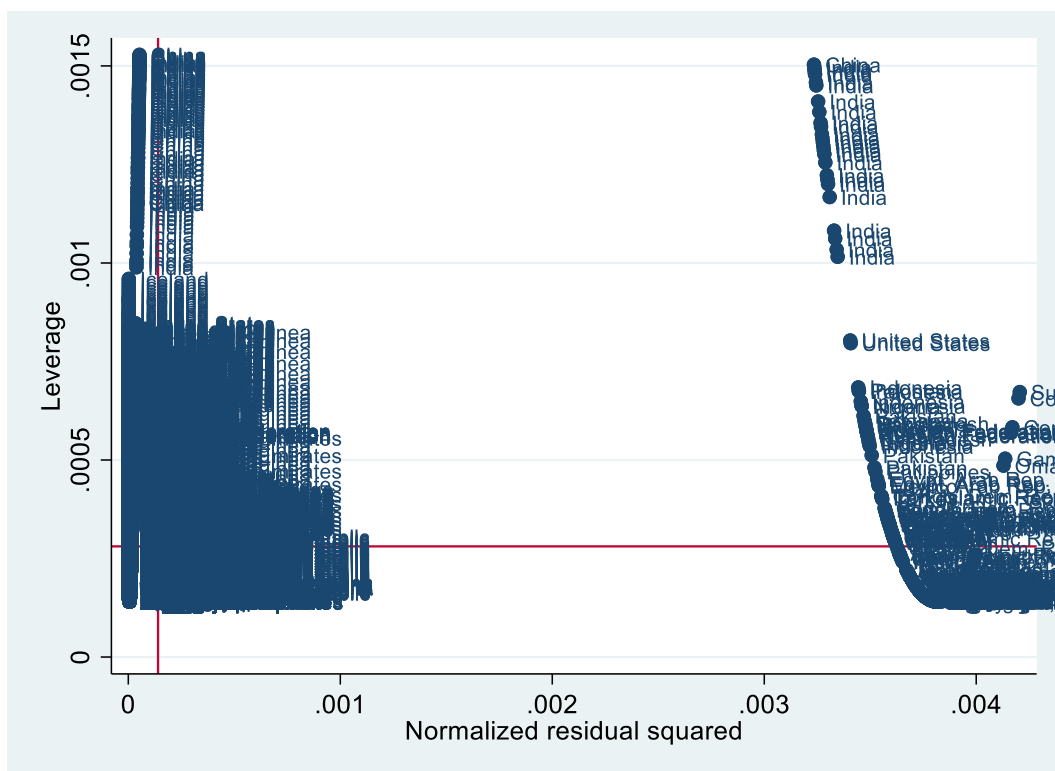
Tabell 4: Korrelasjonsanalyse

3.6.5.3 Ekstremverdier

Det er all grunn til å tro at datagrunnlaget for denne oppgaven kan bestå av en eller flere ekstremverdier. Derfor testes modell 1 for dette, slik at man kan være sikker på at det ikke er noen få observasjoner som bærer hele befolkningsvariabelen. Man kan skille mellom tre typer av ekstremverdier; uteliggere, tunge enheter og betydningsfulle enheter. Uteliggere er kjennetegnet ved store residualer og store avvik mellom predikerte og observerte verdier. Avviket skyldes uvanlig verdi på den avhengige variabelen gitt verdiene på

forklaringsvariablene. Tunge enheter skiller seg ut i form av uvanlige verdier på forklaringsvariablene. Tyngde er bare et problem dersom den tunge enheten i tillegg er en uteligger. Hvis enheten både er tung og uteliggende, kalles verdien for en betydningsfull enhet, og disse kan påvirke koeffisientene i regresjonsanalysene (Midtbø, 2016:116).

De betydningsfulle enhetene for modell 1 er illustrert i et lvr2plot i Figur 8 som viser tyngdeverdiene opp mot de kvadrerte residualene. Enheter med stor tyngde havner øverst på den vertikale Y-aksen. Enhetene langt til høyre i diagrammet er å regne som uteliggere. Betydningsfulle enheter befinner seg i hjørnet øverst til høyre (Midtbø, 2016:117).



Figur 8: lvr2plot over uteliggere og betydningsfulle enheter.

Som en kan se utfra dette plottet kan det virke som at noen observasjoner av Kina og India er å regnes som betydningsfulle enheter i regresjonen mellom befolkningsstørrelse og konflikt. For å undersøke dette nærmere brukes Cooks D-testen, der observasjoner med verdi over $4/N$, er å regnes som betydningsfulle. Denne listen finnes i appendiks. Man kan også se at det er relativt mange uteliggere. I appendiks inkluderes det også en liste over land-år som har standardiserte residualer større enn 3.

Det viser seg å være en del betydningsfulle enheter i datagrunnlaget når man tar utgangspunkt i Cooks D. De to mest betydningsfulle enhetene er Kina 2008 og India 2013. Deretter følger flere observasjoner av Kina og India, og derfor vil det bli kjørt en alternativ regresjonsanalyse uten Kina og India. Samtidig er det verdt å merke seg at disse observasjonene regnes som betydningsfulle ut fra en størrelsesjustert «cut-off». Et annet alternativ er å bruke en absolutt «cut-off» på 1, og da vil det ikke være noen observasjoner som faller inn under kategorien «betydningsfull enhet» (Mehmetoglu & Jakobsen, 2017:156). Det vil derfor gi liten mening å konsekvent droppe disse observasjonene i regresjonsanalysene som danner grunnlag for diskusjonen i denne oppgaven, både ettersom observasjonene er en naturlig del av datagrunnlaget og fordi observasjonene ikke er over den absolutte «cut-off»-punkt på Cooks D-testen (Midtbø, 2016:119).

4. Analyse

I dette kapittelet vil de ulike regresjonsmodellene som har blitt kjørt for å teste hypotesene presenteres og kommenteres. Første del inkluderer modellene som tester hypotesene på hele datasettet. Den neste delen tester resultatene fra de første modellene på datasettet uten Kina og India for alle år, da disse skiller seg ut fra resten med svært store folketall. Modellene har også blitt kjørt uten Kina i 2008 og India i 2013 fordi disse pekte seg ut som de mest ekstreme verdiene, og på datagrunnlaget delt opp i to ulike tidsperioder; 1961-1990 og 1991-2017. Regresjonsanalysene og kommentarer til disse finnes i appendiks.

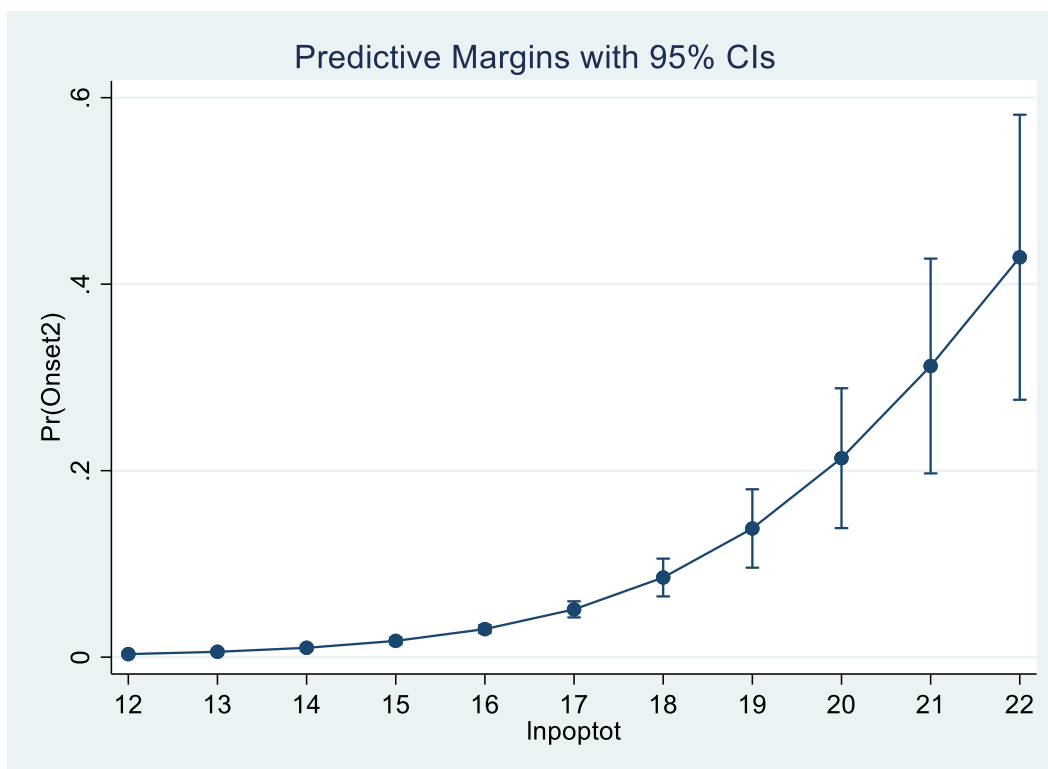
For å teste hypotesene kjøres det 8 regresjonsmodeller, der den første regresjonsmodellen kun inneholder befolkningsstørrelse og konflikt. Deretter følger regresjonsmodellene der hver enkelt variabel blir satt inn sammen med befolkningsstørrelse for å se om de påvirker sammenhengen mellom befolkning og konflikt. Alle modellene kjøres med kontrollvariabelen «fredsår» for å kontrollere for antall år siden siste konflikt. BNP per capita er også inkludert i alle modellene som en kontrollvariabel. Dette er gjort fordi økonomisk tilgang, på lik linje som befolkning, er et robust resultat i Hegre & Sambanis (2006:509) sin studie som sjekker ulike variabler opp mot ulike datasett og operasjonaliseringer av konflikt, i tillegg til at begge variablene på generelt grunnlag er vanlig å kontrollere for. Kontrollvariablene sammen med de 3 _spline-variablene tilhørende fredsårvariabelen vil ikke bli kommentert i analysen, ettersom de kun er med som kontrollvariabler. På generelt grunnlag viser de ulike modellene at BNP per capita som oftest er signifikant med et negativt fortegn, som vil si at høyere BNP per capita gir mindre sannsynlighet for konflikt. Variabelen som måler fredsår er sjeldnere signifikant, og har også et mindre konsekvent resultat, der det finnes både modeller med positivt og negativt fortegn. Dette kan komme av et mulig kurvelineært forhold, da tidligere forskning har funnet resultater som indikerer at sannsynligheten for ny konflikt øker de første årene etter en konflikt, før det gradvis synker (Collier & Sambanis, 2002:5). Collier med flere (2003:83) viser blant annet at et land som akkurat har fått fred har 44% sannsynlighet for ny konflikt innen fem år.

4.1 Modeller kjørt på hele datasettet (1961-2017)

Som tidligere nevnt, er de første regresjonsanalysene kjørt på hele datasettet som omfatter 7123 observasjoner fra årene 1961 til 2017. Modellene viser, som forventet, at befolkning har en positiv effekt på konfliktvariabelen. Det vil si at høye verdier på befolkning går sammen

med høye verdier på konfliktvariabelen. Dette bekrefter tidligere funn, og danner et helt nødvendig grunnlag for studiens videre formål (Hegre & Sambanis, 2006; Sambanis, 2004). Funnet kan man se i Tabell 5 der befolkningsstørrelse har et positivt fortegn og er signifikant på 0,001-nivå i modell 1, 2, 3, 4, 6 og 7, på 0,05-nivå i modell 8 og 8b, og ikke signifikant i modell 5 som inkluderer variabelen for antall unge menn. Utgangspunktet for å kunne si noe om hvorvidt hver enkelt variabel kan forklare sammenhengen mellom befolkningsstørrelse og konflikt er modell 1, der koeffisienten for befolkningsstørrelse er positiv på 0,361. Altså nærmer konfliktvariabelen seg verdien «1» og konflikt for hver gang befolkningen øker. Befolkningsvariabelen bør derfor ha en lavere verdi enn dette i de neste modellene for at de skal kunne forklare sammenhengen mellom befolkning og konflikt.

Et marginsplot av sammenhengen mellom konflikt og befolkningsstørrelse vises i Figur 9. Man kan se at sannsynligheten for konflikt går relativt sakte oppover i starten, før sannsynligheten etter hvert øker mer og mer jo større befolkningen blir.



Figur 9: Marginsplot av sammenhengen mellom konflikt og befolkning i modell 1.

Modell 2 inneholder variabelen som måler totalt areal, og som tester hypotese 1. Denne variabelen endrer befolkningskoeffisienten fra å være 0,361 i modell 1, til å bli 0,272 i modell 2. Befolkningsvariabelen er signifikant på 0,001-nivå, mens arealvariabelen er signifikant på

0,05-nivå. Nedgangen i befolkningskoeffisienten tilsvarer en prosentvis endring på 24,6%. Det vil si at å inkludere totalt areal i modellen gjør at befolkningsvariabelen har en lavere forklaringskraft på konflikt. Altså stjeler arealvariabelen noe av effekten befolkning har på konflikt.

Modell 3 og 4 tester hypotese 2 og 3, og inneholder variablene «relevante grupper» og «ekskluderte grupper». I begge modellene har befolkningsvariabelen en signifikant koeffisient på 0,001-nivå som også er mindre enn i modell 1, henholdsvis 0,284 og 0,345. Endringen er størst i modell tre, der befolkningsvariabelen har en nedgang på 21,3%. I modell 4 er endringen på 4,4%. Variabelen som måler antall relevante grupper er signifikant på 0,05-nivå i modell 3, mens variabelen som måler antall ekskluderte grupper ikke er signifikant i modell 4 med en p-verdi på 0,505. Disse resultatene gir indikasjoner på at antall relevante grupper i et land stjeler mer forklaringskraft fra befolkningsvariabelen enn antall ekskluderte grupper gjør. Med utgangspunkt i litteraturgjennomgangen er det også kjent at antall grupper kan ha et kurvelineært forhold til konflikt, og i neste delkapittel vil dette bli sjekket for å se hvordan dette eventuelt vil påvirke forholdet mellom befolkning og konflikt.

Modell 5 er den første, og eneste modellen i Tabell 5, der befolkningsvariabelen ikke er signifikant. Modell 5 operasjonaliserer hypotese 4 om at en stor befolkning fører til større sjanse for konflikt fordi det er flere unge menn. Begge variablene har høye p-verdier som ikke er i nærheten av å være signifikante. En sannsynlig grunn til dette er den høye korrelasjonen på 0,994 mellom befolkning og unge menn, som praktisk talt betyr at de måler det samme. Dette gjør at det vil være vanskelig å identifisere regresjonskoeffisientene og størrelsene på disse (Midtbø, 2016:128). Det vil derfor være vanskelig å si noe generelt om modell 5, siden den ikke gir noen signifikante resultater. Befolkningskoeffisienten i seg selv viser imidlertid, sett bort ifra at den ikke er signifikant, en nedgang i forklaringskraft sammenliknet med modell 1 på 32,7%. For å sjekke om modellen er signifikant bedre enn modell 1, kjøres en *lr*-test, som brukes for å evaluere forbedring fra en modell til en annen modell (Mehmetoglu & Jakobsen, 2017:212). Denne testen viser en liten bedring i modellen, men med en p-verdi på 0,82, som gjør at dette resultatet heller ikke er signifikant.

I modell 6 testes det om å legge til rural befolkning i modellen minker befolkningsvariabelens forklaringskraft på konflikt, slik hypotese 5 impliserer. Befolkningsvariabelen er signifikant på 0,001-nivå, mens rural befolkning ikke er signifikant med en p-verdi på 0,970. Den

prosentvise nedgangen i befolkningsvariabelen sammenliknet med modell 1 er til stede, men den er svært liten, med 0,8%. Når modellen kjøres gjennom en lrtest viser den heller ikke at modell 6 er bedre enn modell 1. En inkludering av rural befolkning i modell 6 viser seg altså å ikke ha et stort utslag på befolkningskoeffisienten. Modellen sett under ett har heller ikke fått en større forklaringskraft enn modell 1, da justert R^2 er på 0,103 i både modell 1 og 6.

Modell 7 og 8 inneholder variabelen for politisk korrupsjon og et lands totale BNP. Disse operasjonaliserer hypotese 6 om at befolkningsrike land opplever mer konflikt fordi det kan være mye å hente ved å styre landet. I modell 7 er befolkningsvariabelen signifikant på 0,001-nivå, og det samme gjelder politisk korrupsjon. Ved å inkludere korrupsjonsvariabelen i modell 7 øker imidlertid befolkningsvariabelens koeffisient, som vil si at befolkning har en større påvirkning på konflikt når korrupsjon legges til. Denne økningen tilsvarer 3,3%. I praksis vil dette bety at der to land har lik score på korrupsjonsindeksen, og ulik befolkningsstørrelse, vil korrupsjon gi et større utslag på konfliktrisiko i landet med flest folk.

Modell 8, der BNP er inkludert som mulig forklaringsvariabel, har en sterk nedgang i befolkningsvariabelens forklaringskraft på konflikt. Nedgangen er på 34,9%, og representerer dermed den største nedgangen i Tabell 5. Befolkningsvariabelen er signifikant på 0,05-nivå, som er mindre enn for de resterende modellene med signifikant befolkningsvariabel. BNP har en p-verdi på 0,250, som vil si at den ikke er signifikant. Dette resultatet kan også være preget av høy multikollinearitet, ettersom BNP og BNP per capita er sterkt korrelert. På samme måte som i modell 5, kan dette påvirke regresjonskoeffisientene i modellen, og den store nedgangen i forklaringskraft for befolkningsvariabelen bør derfor undersøkes nærmere.

Fordi korrupsjon alene ikke stjeler noe forklaringskraft fra befolkning i modell 7, og BNP stjeler veldig mye i modell 8, har det blitt inkludert en alternativ modell, 8b. I modell 8b er både korrupsjon og BNP lagt til for å sjekke hvordan disse to sammen påvirker befolkningsvariabelens påvirkning på konflikt. Dette er gjort fordi hypotese 7 fordrer at man tester for at både korrupsjon og høy BNP er til stede, slik at det er mulig for de som sitter med makta i et land å ta ut store summer selv. Modell 8b viser en signifikant nedgang i befolkningsvariabelens koeffisient på 19,1% sammenliknet mot modell 1. Befolkning og korrupsjon er signifikant på 0,05-nivå, mens BNP ikke er signifikant.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(8b)
	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2
(ln) befolkning	0.361*** (0.0398)	0.272*** (0.0537)	0.284*** (0.0499)	0.345*** (0.0467)	0.243 (0.529)	0.358*** (0.106)	0.373*** (0.0410)	0.235* (0.117)	0.292* (0.119)
(ln) areal		0.132* (0.0549)							
Relevante grupper			0.0259* (0.0102)						
Ekskluderte grupper				0.00730 (0.0110)					
(ln) unge menn					0.116 (0.518)				
(ln) rural befolkning						0.00357 (0.0957)			
korrupsjon							0.780** (0.293)		0.747* (0.297)
(ln) bnp								0.121 (0.105)	0.0757 (0.105)
fredsår	0.047 (0.0674)	0.041 (0.0675)	0.042 (0.0676)	0.045 (0.0675)	0.048 (0.0675)	0.047 (0.0674)	0.060 (0.0675)	0.063 (0.0685)	0.069 (0.0685)
(ln) bnp p.cap	-0.254*** (0.0498)	-0.267*** (0.0503)	-0.239*** (0.0501)	-0.252*** (0.0498)	-0.256*** (0.0509)	-0.253*** (0.0590)	-0.193*** (0.0563)	-0.365*** (0.109)	-0.266* (0.115)
_cons	-6.845*** (0.795)	-6.950*** (0.795)	-5.799*** (0.898)	-6.606*** (0.874)	-6.409** (2.104)	-6.850*** (0.806)	-8.000*** (0.921)	-6.734*** (0.801)	-7.879*** (0.936)
N	7123	7123	7123	7123	7123	7123	7123	7123	7123
pseudo R ²	0.103	0.106	0.106	0.103	0.103	0.103	0.107	0.104	0.107

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Modellen inneholder også 3_splines som ikke rapporteres

Tabell 5: Modell 1-8, 1961-2017

4.2 Kurvelineært forhold mellom konflikt og etniske grupper

For å sjekke for om et eventuelt kurvelineært forhold mellom konflikt og etniske grupper har betydning for sammenhengen mellom befolkning og konflikt, har det blitt kjørt to modeller (3b+4b) der variablene er opphøyde i annen, slik det er vist i Tabell 6. Variabelen som måler etniske grupper er som sagt allerede signifikant uten andregradsleddet på 0,05-nivå, og i modell 3b er den signifikant på 0,001-nivå. Befolkningsvariabelen er også signifikant. Siden både den opprinnelige variabelen og andregradsleddet er signifikant, og med forskjellig fortegn, tyder dette på et kurvelineært forhold mellom konflikt og relevante grupper, slik også deler av den tidlige litteraturen har pekt på (Ellingsen, 2000:233). I tillegg har

befolkningsvariabelen mistet enda mer forklaringskraft i modell 3b til sammenlikning med den ordinære modell 3. Dette kan tolkes som at antall relevante grupper forklarer en større del av sammenhengen mellom befolkning og konflikt når sammenhengen mellom etniske grupper og konflikt behandles som et kurvelineært forhold. Endringen fra modell 3 til 3b tilsvarer 18,3%, og totalt har inkluderingen av variabelen som måler antall relevante grupper med et andregradsledd i modell 3b minket forklaringskraften til befolkningsstørrelse med 35,7% sammenliknet med koeffisienten i modell 1.

I modell 4 er ikke antall ekskluderte grupper signifikant. Når andregradsleddet blir inkludert i modell 4b ser man at både den opprinnelige variabelen og andregradsleddet blir signifikante på 0,05-nivå, og med ulikt fortegn. I tillegg har koeffisienten til befolkningsvariabelen sunket fra 0,345 til 0,301. Modell 4b styrker dermed hypotese 3, og gir bedre grunnlag for å si noe om hvorvidt ekskluderte grupper har noe å si på forholdet mellom befolkning og konflikt enn hva modell 4 hadde. Koeffisienten til befolkningsvariabelen har sunket med 12,7% fra modell 4 til 4b. Totalt har befolkningsvariabelens forklaringskraft sunket med 16,6% sammenliknet med modell 1.

Antall etniske relevante grupper og ekskluderte grupper viser seg altså å ha et kurvelineært forhold til konflikt, slik litteraturgjennomgangen også viste at det kunne ha (Ellingsen, 2000:233). Samtidig har modell 3b og 4b vist at det kurvelineære forholdet også påvirker befolkningsvariabelens forklaringskraft på konflikt. Når det kontrolleres for kurvelinearitet, synker befolkningsvariabelens forklaringskraft på konflikt i begge modellene.

	(3) onset2	(3b) onset2	(4) onset2	(4b) onset2
(ln) befolkning	0.284*** (0.0499)	0.232*** (0.0518)	0.345*** (0.0467)	0.301*** (0.0495)
Relevante grupper	0.0259* (0.0102)	0.128*** (0.0321)		
Relevante grupper^2		-0.00269** (0.000826)		
Ekskluderte grupper			0.00730 (0.0110)	0.101* (0.0393)
Ekskluderte grupper^2				-0.00254*
fredsår	0.0422 (0.0676)	0.0548 (0.0681)	0.0450 (0.0675)	0.0544 (0.0679)
(ln) bnp p.cap	-0.239*** (0.0501)	-0.203*** (0.0514)	-0.252*** (0.0498)	-0.238*** (0.0501)
_cons	-5.799*** (0.898)	-5.642*** (0.897)	-6.606*** (0.874)	-6.211*** (0.883)
<i>N</i>	7123	7123	7123	7123
pseudo <i>R</i> ²	0.106	0.111	0.103	0.106

Standard errors in parentheses

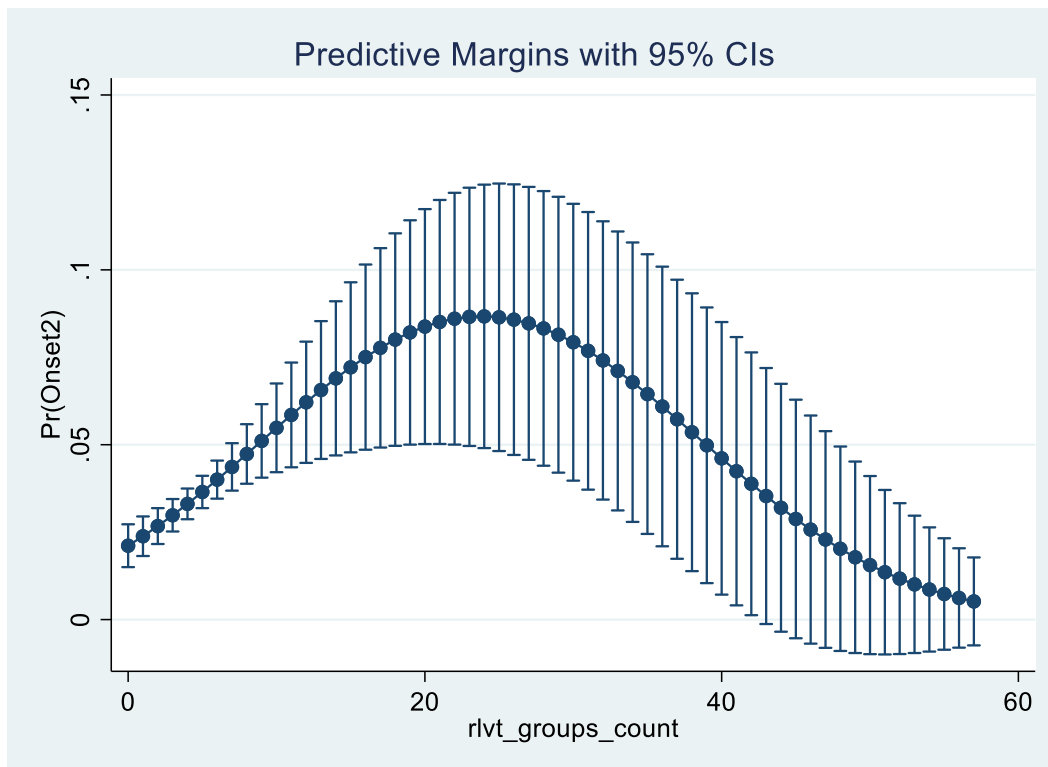
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Modellen inneholder også 3_splines som ikke rapporteres

Tabell 6: Modell 3+4, kurvilinear forhold

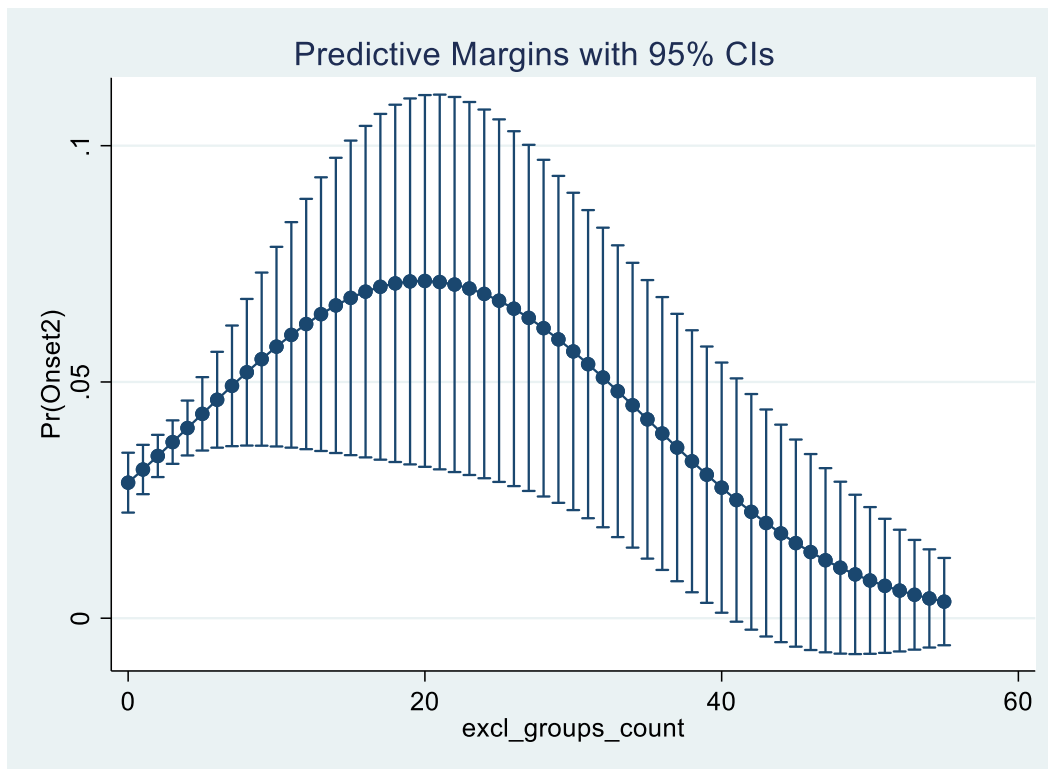
Ved å predikere verdiene til variablene som måler antall relevante og antall ekskluderte grupper med utgangspunkt i modell 3b og 4b, kan man se hvor påvirkningen disse variablene har på konflikt snur. I Figur 10 ser man at modell 3b har størst sannsynlighet for konflikt når det er 24 relevante etniske grupper, før sannsynligheten går ned igjen når alle andre variabler i modellen er målt på gjennomsnittet. Predikasjonsverdiene er signifikante frem til for 42 grupper. For 24 relevante etniske grupper er den isolerte verdien på konfliktvariabelen 0,087.

Figur 11 viser at det tilsvarende punktet for størst isolert sannsynlighet for konflikt er ved 20 ekskluderte grupper, før sannsynligheten synker når alle andre variabler i modellen er målt på gjennomsnittet. Predikasjonen er signifikant frem til for 40 grupper. Den predikerte verdien på konfliktvariabelen er altså høyest ved 20 ekskluderte grupper, og da er den på 0,071. For hypotese 3 betyr dette at sannsynligheten for konflikt øker frem til det er 20 ekskluderte grupper i landet, og deretter er sannsynligheten for konflikt mindre for hvert antall grupper som legges til. Sammenlikner man predikasjonene for de to ulike modellene kan man se at det er antall relevante grupper, og ikke antall ekskluderte grupper, som på generelt grunnlag gir større sannsynlighet for konflikt. At de predikerte verdiene er signifikante frem til henholdsvis 42 og 40 antall grupper, vil si at modellen predikerer godt frem til dette antallet, før de

predikerte resultatene blir noe mer usikre. De to ulike variablene går opp til 57 og 55 grupper, så det er snakk om at modellene predikerer dårlig for de 15 høyeste verdiene. Man kan også se at det kurvelineære forholdet ikke påvirkes av manglende signifikansnivå, da toppene ligger et godt stykke før verdiene som ikke predikerer signifikante verdier. Samtidig er det få observasjoner med flere etniske grupper enn 20. Gjennomsnittet for variablene «relevante grupper» og «ekskluderte grupper» er 3,9 og 2,2, som vil si at vi i praksis likevel kan snakke om en tilnærmet lineær effekt.



Figur 10: Marginsplot som viser sammenhengen mellom konflikt og antall relevante etniske grupper.



Figur 11: Marginsplot som viser sammenhengen mellom konflikt og antall ekskluderte grupper.

4.3 Modeller kjørt på datasettet uten Kina og India

Siden Kina og India har såpass mye større befolkning enn de andre landene, kjøres regresjonene også på et datasett hvor disse landene er kastet ut. Dette gjøres også fordi at de fleste betydningsfulle enhetene er observasjoner på nettopp Kina og India. Modellene som vises i Tabell 7 kan derfor muligens være mer representative for tendensen i verden totalt, da Kina og India skiller seg betydelig ut fra resten av verden.

I Tabell 7 er det noen interessante funn. For modell 10, 11b og 12b er resultatene ganske like, med både signifikant tilleggsvariabel og signifikant befolkningsvariabel som har lavere koeffisient sammenliknet med modell 9. For modell 10 er nedgangen i befolkningsvariabelens styrke på 31,2%. For modell 11b er den 25,4%, og for 12b er den 27,3%. Mens det i Tabell 5 var antall etniske grupper som stjal mest forklaringskraft fra befolkningsvariabelen av disse tre, er det areal som stjeler mest i Tabell 7. Hvis man ser på de ulike modellenes samlede forklaringskraft, er det modellene med relevante og ekskluderte grupper som forklarer mest av variansen i borgerkrigvariabelen med 9,1%, mens det er inkluderingen av areal som variabel som stjeler mest av befolkningsvariabelens forklaringskraft på konflikt.

Modell 14 og 15, med rural befolkning og korrupsjon, har derimot større endringer når Kina og India ikke er inkludert. I modell 14 er befolkningsvariabelen fortsatt signifikant, men den har nå fått en høyere befolkningskoeffisient enn i modell 9, som vil si at å inkludere rural befolkning øker befolkningsvariabelens forklaringskraft på konflikt. For modell 15 har det motsatte skjedd. Her har befolkningsvariabelen nå blitt lavere enn hva den er i modell 9, som vil si at inkludering av korrupsjonsvariabelen gjør at befolkning har mindre effekt på sannsynlighet for konflikt enn tidligere. For modell 16 er befolkningskoeffisienten fortsatt veldig lav sammenliknet med modell 9, men nå er den ikke signifikant lenger. Det er heller ikke BNP. Modell 16b gir heller ikke signifikant befolkningsvariabel.

Det viktigste man kan lese ut av disse resultatene er at areal blir viktigere som forklaringsvariabel på sammenhengen mellom befolkning og konflikt når Kina og India ikke er med. Det samme gjelder ekskluderte grupper, som viser seg å være viktigere enn relevante grupper sammenlignet med Tabell 5. Rural befolkning stjeler ikke lenger forklaringskraft fra befolkningsvariabelens påvirkning på konflikt, som vil si at observasjoner for Kina og India teller mye inn på samspillet som man kunne se i Tabell 5. Dette er et viktig resultat. At modell 16 ikke har en signifikant befolkningsvariabel er også et viktig funn fra disse modellene, ettersom det forteller oss noe om hvor robust det noe overraskende funnet om at BNP stjeler mest forklaringskraft fra befolkning i Tabell 5 er. Et ikke-signifikant resultat i en modell uten Kina og India kan tolkes som at det tidligere resultatet avhenger mye av nettopp Kina og Indias observasjoner på BNP.

	(9)	(10)	(11)	(11b)	(12)	(12b)	(13)	(14)	(15)	(16)	(16b)
	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2
(ln) Befolkning	0.362*** (0.0517)	0.249*** (0.0630)	0.298*** (0.0538)	0.270*** (0.0548)	0.303*** (0.0536)	0.263*** (0.0554)	0.375 (0.544)	0.371*** (0.107)	0.330*** (0.0525)	0.223 (0.127)	0.237 (0.126)
(ln) areal		0.170** (0.0558)									
Relevante grupper			0.0506*** (0.0112)	0.113*** (0.0323)							
Relevante grupper^2				-0.00151* (0.000754)							
Ekskluderte grupper					0.0506*** (0.0114)	0.136*** (0.0371)					
Ekskluderte grupper^2						-0.00211* (0.000901)					
(ln) unge menn							-0.0126 (0.533)				
(ln) rural befolkning								-0.00862 (0.0896)			
Korrupsjon									1.029** (0.336)		0.997** (0.339)
(ln) BNP										0.129 (0.108)	0.0883 (0.108)
Fredsår	0.126 (0.0701)	0.123 (0.0703)	0.148* (0.0709)	0.149* (0.0710)	0.152* (0.0711)	0.161* (0.0714)	0.126 (0.0702)	0.126 (0.0701)	0.138* (0.0703)	0.143* (0.0714)	0.149* (0.0714)
(Ln)bnp p.cap	-0.268*** (0.0530)	-0.287*** (0.0541)	-0.280*** (0.0537)	-0.260*** (0.0544)	-0.291*** (0.0540)	-0.279*** (0.0539)	-0.267*** (0.0541)	-0.271*** (0.0615)	-0.160* (0.0650)	-0.383*** (0.111)	-0.243* (0.120)
_cons	-6.891*** (0.917)	-7.059*** (0.923)	-6.043*** (0.943)	-5.980*** (0.939)	-5.952*** (0.951)	-5.598*** (0.956)	-6.937** (2.166)	-6.880*** (0.924)	-7.857*** (0.988)	-6.747*** (0.926)	-7.728*** (1.001)
<i>N</i>	7009	7009	7009	7009	7009	7009	7009	7009	7009	7009	7009
pseudo <i>R</i> ²	0.080	0.085	0.089	0.091	0.088	0.091	0.080	0.080	0.085	0.081	0.086

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Modellen inneholder også 3_splines som ikke rapporteres

Tabell 7: Modell 9-16, uten Kina og India

4.4 Samlet forklaringskraft

Selv om oppgavens hovedformål har vært å avdekke hvilke aspekter ved befolkningsrike land som kan være med på å forklare sammenhengen mellom befolkningsstørrelse og konflikt, vil det også være interessant å kort kommentere hvor mye av befolkningsvariabelens forklaringskraft som faller vekk når variablene «areal», «etnisk relevante grupper» og «ekskluderte etniske grupper» legges inn i en enkelt modell. Disse variablene er plukket ut på grunnlag av at de gir de klareste resultatene i modellene som er kjørt. Modell 17 viser en befolkningskoeffisient som er betydelig lavere sammenliknet med modell 1. Den har gått fra å være 0,361 til å bli 0,135, som tilsvarer en prosentvis nedgang på 62%. Over halvparten av befolkningsvariabelens forklaringskraft forsvinner altså når de tre variablene som står for de mest robuste resultatene inkluderes i en felles modell i Tabell 8. Befolkningsvariabelen er også signifikant, til tross for at modellen inneholder flere ulike variabler. R^2 viser også en forbedring i modellens totale forklaringskraft på variansen i den avhengige variabelen, sammenliknet med de andre modellene som er kjørt på det samme datagrunnlaget i Tabell 5.

	(17) onset2
(ln) befolkning	0.135* (0.0625)
(ln) areal	0.127* (0.0565)
Relevante grupper	0.0281 (0.0632)
Relevante grupper^2	0.00253 (0.00253)
Ekskluderte grupper	0.0543 (0.0637)
Ekskluderte grupper^2	-0.00451 (0.00253)
fredsår	0.0633 (0.0689)
(ln) BNP p.cap	-0.214*** (0.0527)
_cons	-5.367*** (0.956)
<i>N</i>	7123
pseudo <i>R</i> ²	0.115

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Modellen inneholder også 3_splines som ikke rapporteres

Tabell 8: Modell 17, fullmodell

4.5 Sammenfattende analyse

Etter å ha kjørt samtlige modeller sitter man igjen med noen resultater som viser seg å være mer gjeldende enn andre. De variablene som viser seg å gi omtrentlige like resultater i alle modellene er areal, antall relevante etniske grupper og antall ekskluderte grupper. Disse variablene har påvirket befolkningsvariabelens koeffisient på en slik måte at man ser en nedgang når man sammenlikner med modellen der bare befolkning er lagt til. For modellen med antall unge menn viser alle modellene som er kjørt et fravær av signifikante koeffisienter på både befolkningsvariabelen og variabelen som måler unge menn.

Sett under ett kan man ikke sammenlikne alle modellene som har blitt kjørt på grunn av at utvalget ikke er det samme. Derfor vil de ulike variablenes påvirkning sammenliknes med

utgangspunkt i Tabell 5, som inneholder modellene som er kjørt på hele datagrunnlaget. Det vil bli tatt utgangspunkt i den prosentvise endringen i befolkningsvariabelen i sammenlikningen opp mot modell 1, der den står alene med kontrollvariablene for fredsår og BNP per capita. Modellene i Tabell 9, 10 og 11, som finnes i appendiks, vil bli brukt som supplement for å kunne si noe om robustheten til de ulike resultatene der det er behov for det.

Variabelen som stjeler mest forklaringskraft fra befolkningsvariabelen i Tabell 5 og 6 er antall relevante grupper² med 35,7% nedgang i forklaringskraft. Etter dette følger BNP med en nedgang på 34,9%, før areal med en nedgang på 24,6%. Modellen som inkluderer både BNP og korrupsjon, har en prosentvis nedgang på 19,1%. Deretter kommer ekskluderte grupper med en nedgang på 16,6%, før rural befolkning med en nedgang på 0,8%. Korrupsjon har en økning på 3,3%, mens unge menn har en nedgang på 32,7%, uten at resultatet er signifikant.

Med utgangspunkt i dette kan man gå tilbake til hypotesene og vurdere om de skal forkastes eller beholdes. For hypotese 1, «*Land med stor befolkning opplever mer borgerkrig fordi de også har store territorium som kan gjøre det utfordrende for staten å ha kontroll på hele området sitt, og som dermed gjør det lettere for befolkningen å gjøre opprør*», finnes det belegg for å si at stort areal kan være en forklaringsfaktor for at befolkningsrike land opplever mer konflikt enn mindre land. Både befolkning og areal har begge signifikante p-verdier, i tillegg til at justert R² er høyere for modell 2 enn for modell 1, som gir oss gode indikasjoner på at resultatene er pålitelige. Disse funnene bekrefter også tidligere funn om at areal har en påvirkning på konfliktrisiko, slik blant annet Fearon & Laitin (2003), Collier & Hoeffler (2004) og Buhaug (2006) også finner. At hypotese 1 beholdes betyr også at forholdet mellom befolkningsstørrelse og konflikt kan skyldes utfordringene som følger med et stort areal, og er på så måte ny innsikt i hvordan areal, befolkningsstørrelse og konflikt henger sammen. Faktorer som spiller inn på dette forholdet kan blant annet være at staten kan ha mindre makt og kontroll i periferien, og at potensielle opprørsgrupper har lettere for å vokse langt unna hovedstaten (Buhaug, 2006; Butler & Gates, 2009; Collier & Hoeffler, 2004; Fearon & Laitin, 2003; Linke m.fl., 2017).

Hypotese 2, «*Land med stor befolkning opplever mer borgerkrig fordi det er flere etniske grupper innad i befolkningen som kan være misfornøyde og gå til opprør*», beholdes og bekreftes i Tabell 5 og 6. Befolkningsvariabelen har en betydelig mindre koeffisient når antall etniske grupper inkluderes, og begge variablene er signifikante. Det kurvelineære forholdet

mellom konflikt og etniske grupper viser oss at antall grupper gir økende risiko frem til 24 grupper, før det synker nedover igjen. Den forklarte variansen for modell 3b er 11,1% og er den høyeste verdien sammenliknet med alle modellene. Dette er også med på å støtte opp om at hypotesen bør beholdes. Sammenhengen mellom befolkning og konflikt kan, med utgangspunkt i hypotese 2, forklares utfra at etnisitet samler folket i ulike ikke-diskuterbare grupper som samfunnet fordeler seg utfra (Reynal-Querol, 2002; Denny & Walter, 2014). Dette kan tyde på at en samling rundt etnisitet gjør det lettere å mobilisere fordi opprørsgruppene samles på bakgrunn av felles språk og kultur (Bormann m.fl., 2017). Funnet i disse modellene sammenfaller med tidligere forskning, som for eksempel Ellingsen (2000) sine funn om at antall grupper innad i et land er relatert til konflikt i en omvendt U-kurve, i stedet for at jo høyere antall grupper jo høyere er risikoen. Funnene faller ikke like godt sammen med blant annet Cunningham & Weidmann (2010) som hevdet at det er marginalisering eller ekskludering av grupper som har noe å si for konfliktrisiko, og ikke antallet etniske grupper innad i et land, da resultatene fra denne analysen viser at både antall etniske grupper og antall diskriminerte grupper har en signifikant påvirkning på konflikt. I denne studien viser det seg både signifikante resultater for et kurvelineært og et lineært forhold mellom konflikt og antall etniske grupper, men det er altså det kurvelineære forholdet som påvirker sammenhengen mellom befolkning og konflikt i størst grad, og som i så måte er et nytt funn.

Hypotese 3, «Land med stor befolkning opplever mer borgerkrig fordi det er flere ekskluderte etniske grupper innad i befolkningen som kan være misfornøyde og gå til opprør», beholdes og bekreftes med utgangspunkt i Tabell 6. Resultatet er det samme i alle modellene, med unntak av for årstallene 1961-1990, der inkluderingen gjør at befolkningsvariabelen har en større forklaringskraft på konflikt. I Tabell 5 er ikke antall ekskluderte grupper signifikant, men når det inkluderes et andregradsledd endrer dette seg. Begge variablene i modell 4b er signifikante, og befolkningsvariabelen har en lavere koeffisient sammenliknet med modell 1. Det kurvelineære forholdet viser også her at antall diskriminerte grupper øker sannsynligheten for konflikt frem til 21 grupper, før det synker nedover. I modell 4 er ikke justert R^2 høyere enn for modell 1, men inkluderingen av andregradsleddet i modell 4b gjør at også denne modellen har en høyere forklaringskraft enn modell 1, som gir støtte for at også hypotese 3 bør beholdes. Det vil si at også antall ekskluderte grupper påvirker sammenhengen mellom befolkning og konflikt, men her kommer det tydeligere frem at det er det kurvelineære forholdet som har en signifikant påvirkning. Funnene er på så måte i tråd med

Cederman med flere (2013) og Bormann med flere (2017), som også viser at antall ekskluderte grupper har en påvirkning på konfliktrisiko. Analysen bidrar også med ny innsikt om temaet ved å vise hvordan både relevante og diskriminerte grupper påvirker sammenhengen mellom konflikt og befolkningsstørrelse.

Hypotese 4, «*Land med stor befolkning opplever mer borgerkrig fordi det består av flere unge menn som anser det som mer lønnsomt å delta i voldelig konflikt enn å bevare status quo*», kan ikke bekreftes med utgangspunkt i regresjonsanalysene, da analysene ikke gir noen klare resultater. Siden verken befolkningsvariabelen eller unge menn er signifikant, vet vi heller ikke om resultatene i modellene er til å stole på. At befolkningsstørrelse og antall unge menn korrelerer så mye som de gjør, gjør at det er vanskelig å tolke p-verdiene og koeffisientene. For å kunne svare på hypotese 4 trengs det alternative måter å måle et lands antall unge menn.

For de resterende hypotesene er det ikke et like klart svar. Når rural befolkning, korrupsjon og BNP inkluderes i modellene, gir de ulike resultater avhengig av hvilket datagrunnlag de testes på. Tar man hele datagrunnlaget som utgangspunkt viser modellene at rural befolkning stjeler litt av befolkningsvariabelens forklaringskraft på konflikt, BNP stjeler mye, mens korrupsjon snarere øker befolkningsvariabelens koeffisient. Deretter gir hver enkelt modell ulike svar i de resterende tabellene. Hypotesene vil gjennomgås nedenfor.

Hypotese 5, «*Land med stor befolkning opplever mer borgerkrig fordi det består av flere mennesker som bor i rurale områder som anser det som mer lønnsomt å delta i voldelig konflikt enn å bevare status quo*», kan til en viss grad beholdes med utgangspunkt i modell 6 som har blitt kjørt. Det er på svakere grunnlag enn for de andre hypotesene, med tanke på at rural befolkning ikke har en signifikant koeffisient. Befolkningsvariabelen er derimot signifikant, og dette kan tolkes i retning av at antall folk som bor på bygda er med og påvirke forholdet mellom befolkning og konflikt. Dette kan blant annet bety at man i fremtiden kan forvente mindre konflikt, med tanke på at verden stadig blir mer urbanisert, og færre bor i rurale områder uten lik tilgang på goder som i byene (Buhaug, 2006). Endringen i befolkningskoeffisienten er imidlertid ganske liten, som gir oss implikasjoner på at antallet som bor i rurale områder trolig ikke har en stor påvirkning på forholdet mellom befolkning og konflikt, om det i det heletatt har det. Derfor brukes resultatene fra de andre tabellene for å se etter et klarere svar.

For rural befolkning gir fraværet av Kina og India i modell 14 en signifikant befolkningskoeffisient som er høyere enn hva befolkning i utgangspunktet er i modell 9. Deretter følger modell 31 for årstallene 1961-1990 hvor befolkningsvariabelen har mistet mye forklaringskraft, men der koeffisienten ikke er signifikant. Til slutt, i modell 39 for årstallene 1991-2017, er befolkningsvariabelen igjen høyere enn hva den er i modellen med bare befolkningsvariabelen, og dette resultatet er også signifikant. Dette gir oss implikasjoner på at funnet i den første modellen ikke er et særlig robust resultat når det kommer til spørsmålet om antall mennesker som bor i rurale områder kan forklare sammenhengen mellom befolkning og konflikt. Det gir oss imidlertid heller ikke et klart svar på at rural befolkning ikke påvirker sammenhengen mellom befolkning og konflikt.

Hypotese 6, «*Land med stor befolkning opplever mer borgerkrig fordi det er mer å tjene på å styre landet på grunn av korrupsjon og tilgang på kapital*», har også et noe uklart resultat med utgangspunkt i Tabell 5. Modell 7 med korrupsjon inkludert, viser en befolkningskoeffisient som er høyere enn i modell 1. Dette tyder på at korrupsjon ikke kan forklare sammenhengen mellom befolkning og konflikt alene, men korrupsjon i seg selv øker sannsynligheten for konflikt, som sammenfaller med resultatene fra blant annet Mcferson (2009) og Le Billion (2003). BNP derimot, står for en stor nedgang i befolkningskoeffisienten, men resultatet er ikke signifikant, og modellens R^2 er bare 0,001 høyere enn for modell 1. Disse resultatene tilsier at høy BNP ikke har noen signifikant påvirkning på sannsynlighet for konflikt, slik vi kunne forventet å se med utgangspunkt i for eksempel Berman med flere (2011). Siden hypotese 6 er avhengig av at det både er korrupsjon og høy BNP til stede for å kunne beholdes, kjøres det en modell 8b, for å sjekke hvordan modellen vil se ut hvis man inkluderer begge variablene i en modell. I modell 8b er justert R^2 bedre, og befolkningsvariabelen har en nedgang på 19,1%. Dette kan tyde på at tilstedeværelse av både korrupsjon og høy BNP gjør at opprørerne i landet er villige til å gjøre mer for å sitte med makta. Siden resultatene fra Tabell 5 gir noe uklare resultater, brukes de andre modellene kjørt på deler av datagrunnlaget for å sjekke om det er mulig å få noen klarere implikasjoner.

BNP er som sagt den variabelen i Tabell 5 som stjeler mest forklaringskraft fra befolkning. I de resterende modellene gir en inkludering av BNP derimot ingen klare implikasjoner. I modellen 25 uten Kina og India og modell 33 for årstallene 1961-1990 har

befolkningsvariabelen en mindre koeffisient sammenliknet med modellene med bare befolkning. Her er imidlertid verken befolkning eller BNP signifikante, som gjør det vanskelig å stole på disse resultatene. I modell 41 som er kjørt for årstallene 1991-2017 har inkludering av BNP i tillegg gjort befolkningsvariabelens koeffisient større enn hva den er i modellen med bare befolkning, og dette resultatet er signifikant. De samme resultatene gjelder også for de alternative modellene der korrupsjon og BNP er satt sammen. For modell 25b er imidlertid befolkningsvariabelen signifikant og mindre enn for modell 18. Det finnes dermed resultater som taler både for og imot at hypotese 6 skal beholdes. Ettersom det også her er et signifikant resultat på hele datasettet, velger jeg å ikke forkaste hypotese 6, på likt grunnlag som for hypotese 5. Resultatene er mindre sikre enn for mange av de andre hypotesene.

Tar man utgangspunkt i denne sammenfattende analysen, står man igjen med tre hypoteser med sikrere resultater sammenliknet med de resterende hypotesene. Dette er hypotesene som omfatter areal, antall etnisk relevante og etnisk ekskluderte grupper. Selv om disse hypotesene bekreftes, betyr det ikke at det alltid vil være tilfelle i den virkelige verden. Befolkningsrike land vil for eksempel ikke alltid også være store i areal, på samme måte som at det ikke alltid vil være mange etniske grupper i land med mange mennesker. Bangladesh er et eksempel på en case som ikke passer like godt inn i modellene denne oppgaven har forespeilet. Landet består av et lite areal, og et relativt gjennomsnittlig antall etniske grupper (rundt 4 relevante, 3 ekskluderte), og har verdens 7. største befolkning. I løpet av perioden 1961-2017 har det vært tre nye konfliktutbrudd. Det samme gjelder for Nigeria, som har et lite areal, litt over gjennomsnittlig antall etniske grupper (rundt 6 relevante, 4 ekskluderte), og er verdens 6. største befolkning. Her har det vært 5 konfliktutbrudd i perioden. Hvis man så antar at disse landene opplever konflikt blant annet på grunn av det store folketallet, vil ikke areal, antall etnisk relevante eller ekskluderte grupper være like relevante forklaringsfaktorer for denne sammenhengen rent empirisk sammenliknet med hva modellene ellers forespeiler at de gjør. På så måte fungerer disse to casene som gode eksempler på at det også finnes andre aspekter som påvirker sammenhengen mellom befolkning og konflikt som disse regresjonsanalysene ikke har klart å fange opp.

India og Indonesia vil på den andre siden representere caser som passer godt inn i modellene som er forespeilet. Begge landene er svært folkerike som nummer 2 og 4 i verden, og har mange etniske grupper. India har flest av de to med rundt 19 relevante og 6 ekskluderte og Indonesia med rundt 13 relevante og 4 ekskluderte grupper. India har verdens 7. største areal,

mens Indonesia er nummer 14. I perioden har India 25 nye utbrudd av konflikt, mens Indonesia har 8 utbrudd. Disse casene stemmer derfor godt overens med resultatene fra regresjonsanalysene, da vi kan se at både et stort folketall, et stort areal og et stort antall etniske grupper er til stede på samme tid i en periode med flere utbrudd i både India og Indonesia.

Det vil også være interessant å diskutere hvordan areal, etniske grupper og ekskluderte grupper forholder seg til befolkningsstørrelse. Selv om befolkning og disse tre variablene korrelerer, er det ikke nødvendigvis slik at det er en enkelt kausal sammenheng mellom dem. Det er for eksempel logisk å tenke at et stort areal fører til et stort folketall, snarere enn motsatt. Dette fordi land som er store i areal også har større plass. At et befolkningsrikt land fører til større areal er selvfølgelig også mulig, men da vil det kanskje være snakk om erobring av land og ekspansjon som årsak til endringen i størrelse, som ikke er et veldig vanlig fenomen i dag. Når det kommer til etniske grupper er det mest nærliggende å tenke at det er befolkningsstørrelsen som har ført til flere antall grupper, og ikke flere antall grupper som har ført til flere folk. Et befolkningsrikt land vil naturlig nok inneholde et større totalt antall mennesker, og det er da en større sannsynlighet for at de tilhører forskjellige grupper. Det er vanskeligere å tenke seg en motsatt årsakssammenheng i dette tilfellet.

Selv om det er mest naturlig å tro at areal fører til stor befolkning, og at stor befolkning fører til et større antall etniske grupper, er det altså vanskelig å si noe helt konkret om dette. Det vil ikke alltid være slik at store areal fører til store befolkninger, som blant annet Kazakhstan er et godt eksempel på. Spørsmålet blir derfor om det er et stort areal som fører til en stor befolkning som igjen fører til større sannsynlighet for konflikt, og om det er en stor befolkning som fører til et stort antall grupper som fører til konflikt. Som oppgavens teorigjennomgang viser, er både areal og antall grupper assosiert med konfliktrisiko, men spørsmålet som blir stående igjen er hvordan dette henger sammen, og hva som påvirker hva, og på hvilken måte. Kanskje er det grunn til å tro at et lands areal er den egentlige synderen for konfliktrisiko, og at å kontrollere for areal vil være like relevant som å kontrollere for befolkning i fremtiden?

5. Konklusjon

Denne oppgavens hovedmålsetning har vært å undersøke hva som kan ligge til grunn for sammenhengen mellom befolkningsstørrelse og konflikt. Forskningsspørsmålet «*Hvorfor opplever befolkningsrike land oftere utbrudd av intern konflikt enn andre?*» har fungert som utgangspunkt for oppgaven. Sammenhengen mellom befolkning og konflikt er kjent for å være robust, men den har blitt tatt for gitt uten å ha fått den forskningen og tiden det fortjener. Gjennom en rekke regresjonsanalyser har derfor seks ulike hypoteser om dette forholdet blitt testet. Det har blitt undersøkt hvorvidt de ulike operasjonaliseringene av disse hypotesene har stjålet forklaringskraft fra befolkningsvariabelen når de har blitt inkludert i analysene. Som analysekapittelet har vist, er det flere av hypotesene som er vanskelig å forkaste på grunnlag av resultatene de har gitt i de ulike regresjonsanalysene. Hypotesene som med størst sikkerhet kan beholdes, er hypotese 1, 2 og 3. Disse måler et lands totale areal, et lands totale antall etnisk relevante grupper og et lands totale antall etnisk ekskluderte grupper sin påvirkningsevne på befolkning som forklaringsfaktor for konflikt. Det vil si at befolkning får en betydelig svakere forklaringskraft på konflikt når disse faktorene kontrolleres for.

Et lands totale areal er en faktor som sjeldent endrer seg noe særlig over tid. Om arealasppektet ved et land var noe som kunne blitt endret, kunne kanskje befolkningsrike land også opplevd mindre konflikt. Antall ekskluderte og relevante grupper er derimot variabler som kan endre seg noe mer fra år til år. Funnene som er relatert til disse variablene viser at det vil være en fordel for befolkningsrike land å inkludere de ulike etniske gruppene tilhørende landet i det offentlige liv, og gi dem de samme mulighetene som resten av befolkningen. Når disse gruppene er inkludert, gir befolkningsstørrelse et vesentlig mindre utslag som risikofaktor for konflikt. Samtidig viser analysene også at det vil være vesentlig mindre konflikt i land der de fleste gruppene er ekskludert. Dette kan blant annet være fordi det er en gruppe som sitter med mye makt som gjør det usannsynlig for de mindre gruppene å kunne vinne et eventuelt opprør.

Den samme tendensen ser man på antall relevante etniske grupper. Et befolkningsrikt land med enten veldig mange eller veldig få relevante etniske grupper, vil ha mindre sannsynlighet for konflikt. Disse funnene tyder på at det vil kunne være en fordel for befolkningsrike land å prøve å finne representanter for enhver etnisk gruppe som kan være med i landets politikk.

Slike resultater gir implikasjoner på at disse aspektene ved land kan forklare sammenhengen mellom befolkningsstørrelse og konflikt. Videre trengs det å undersøkes hvordan disse faktorene påvirker denne sammenhengen. En av svakhetene til dette prosjektet er at de ulike aspektene er behandlet noe overfladisk og på generelt grunnlag. Derfor står det igjen spørsmål som for eksempel hvordan areal påvirker sammenhengen mellom befolkning og konflikt. En smalere tilnærming og operasjonalisering av de ulike variablene som denne analysen har funnet resultater for, bør gi oss en bredere forståelse av hvordan denne sammenhengen er. For videre forskning kunne det vært interessant å for eksempel inkludere mål på avstand fra viktige holdepunkter i landet for å undersøke hypotese 1 nærmere. For hypotese 2 og 3 vil det være interessant å inkludere størrelse på de ulike gruppene, og kanskje også andre aspekter som språk og religion. Som tidligere diskutert vil det også være viktig med en dypere forståelse av den kausale sammenhengen mellom disse variablene, slik at man kan si noe mer konkret om årsaksforholdet. Selv om det er å forvente at areal påvirker befolkningsstørrelse og befolkningsstørrelse påvirker antall etniske relevante og ekskluderte grupper, kan dette med fordel drøftes ytterligere.

Målet for dette prosjektet har ikke vært å finne en enkelt variabel som kunne forklare sammenhengen mellom befolkningsstørrelse og konflikt, men heller vise til flere mulige aspekter som alle kan være med og påvirke denne sammenhengen. Det viktigste prosjektet har bidratt med, er å avklare at areal, antall etniske grupper og antall ekskluderte grupper alle påvirker sammenhengen mellom befolkning og konflikt. Samtidig har også oppgaven bekreftet tidligere funn som knytter de ulike aspektene til konflikt. Studien viser dermed hvilke aspekter det kan være nyttig å gå nærmere inn på, og har på så måte lagt et grunnlag for hva som bør undersøkes videre. Resultatene har gitt en dypere innsikt i et spørsmål som foreløpig har få andre svar, og de kan derfor fungere utmerket som utgangspunkt for videre analyser. Som modell 17 viser, har denne oppgaven klart å finne frem til aspekter som mer enn halverer befolkningsvariabelens forklaringskraft på konflikt. Det vil også si at halvparten står igjen, og at det er rom for videre forskning på dette feltet. Hva skal til for å forklare den resterende forklaringskraften til befolkning?

6. Litteraturliste

Acemoglu, D., Fergusson, L. & Johnson, S. (2019). Population and Conflict. *Review of economic studies*, 0:1-40.

Alesina, A. & Spolaore, E. (2003). *The size of nations*. Cambridge, MA: MIT press.

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Bao, C. & Fang, C-L. (2007). Water resources constraint force on urbanization in water deficient regions: A case study of the Hexi Corridor, arid area of NW China. *Ecological Economics*, 62(3-4):508-517.

Beber, B. & Blattman, C. (2013). The logic of child soldiering and coercion. *International Organization*, 67:65-104.

Beck, N., Katz, J.N. & Tucker, R. (1998). "Taking Time Seriously in Binary Time-Series Cross-Section Analysis." *American Journal of Political Science*, 42(4):1260-1288.

Berman, E., Callen, M., Felter, J. & Shapiro, J. (2011). Do working men rebel? Insurgency and unemployment in Afghanistan, Iraq and the Phillipines. *Journal of conflict resolution*, 55(4):496-528

Blimes, R. (2006). The indirect effect of ethnic heterogeneity on the likelihood of civil war onset. *Journal of conflict resolution*, 50(4):536-547.

Boix, C. (2003). *Democracy and Distribution*. Cambridge: Cambridge University Press.

Bormann, N-C., Cederman, L-E. & Vogt, M. (2017). Language, religion and ethnic civil war. *Journal of conflict resolution*, 61(4):744-771.

Boulding, K. (1962). *Conflict and defense*. New York: Harper and Row.

Bravo-Ortega, C. & de Gregorio, J. (2005). The relative richness of the poor? Natural resources, human capital, and economic growth. Policy research working paper series 3484, The world bank.

Buhaug, H. (2006). Relative capability and rebel objective in civil war. *Journal of Peace Research*, 43(6):691-708.

Buhaug, H. (2010). Dude, where is my conflict? *Conflict Management and Peace Science*, 27(2):107-128.

Buhaug, H. & Gates, S. (2002). The Geography of Civil War. *Journal of Peace Research*, 39(4):417-433.

Buhaug, H. & Rød, J.K. (2006). Local determinants of african civil wars, 1970-2001. *Political geography*, 25(3):315-335.

Butler, C. & Gates, S. (2009). Assymetry, Parity, and (Civil) war: Can International theories of Power help us understand civil war? *International interactions* 35(3):330-340.

Carter, D., Shaver, A. & Wright, A. (2019). Places to hide: terrain, ethnicity and civil conflict. *The journal of politics*, 81(4):1446-1465.

Cederman, L-E., Gleditsch K.S. & Wucherpfennig, J. (2017). Predicting the decline of ethnic civil war: Was Gurr right and for the right reasons? *Journal of peace research*, 54(2):262-274.

Cederman, L-E., Gleditsch, K.S. & Buhaug, H. (2013). *Inequality, Grievances, and Civil War*. New York: Cambridge university press.

Cederman, L., Wimmer, A., & Min, B. (2010). Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics*, 62(1):87-119.

Chomitz, K., Buys, P. & Thomas, T.S. (2005). *Quantifying the rural-urban gradient in Latin America and the Caribbean (English)*. Policy, Research working paper; no. WPS 3634. Washington, DC: World Bank,

Collier, P., Elliott, V.L., Hegre, H., Hoeffler, A., Reynal-Querol, M. & Sambanis, N. (2003). Breaking the conflict trap. Civil War and development policy. A World Bank policy research working paper. Washington, DC: World Bank & Oxford University Press.

Collier, P. & Hoeffler, A. (1998). On economic causes of civil war. *Oxford economic papers*, 50(4):563-573.

Collier, P. & Hoeffler, A. (2004). Greed and Grievance in Civil war. *Oxford Economic Papers*, 56:663-695.

Collier, P. & Sambanis, N. (2002). Understanding Civil War: A new agenda. *Journal of conflict resolution*, 46(1):3-12.

Coppedge, M., Gerring, J., Knutsen, C.H., Lindberg, S.I., Teorell, J., Altman, D., Bernhard, M., Fish, M.S., Glynn, A., Hicken, A., Lührmann, A., Marquardt, K.L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S-E., Staton, J., Wilson, S., Cornell, A., Gastaldi, L., Gjerløw, H., Ilchenko, N., Krusell, J., Maxwell, L., Mechkova, V., Medzihorsky, J., Pernes, J., Von Römer, J., Stepanova, N., Sundström, A., Tzelgov, E., Wang, J-T., Wig, T. & Ziblatt, D. (2019A). «V-Dem [Country-Year/Country-Date] Dataset v9», Varieties of Democracy (V-Dem) Project.

Coppedge, M., Gerring, J., Knutsen, C.H., Lindberg, S.I., Teorell, J., Altman, D., Bernhard, M., Fish, M.S., Glynn, A., Hicken, A., Lührmann, A., Marquardt, K.L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S-E., Staton, J., Wilson, S., Cornell, A., Gastaldi, L., Gjerløw, H., Ilchenko, N., Krusell, J., Maxwell, L., Mechkova, V., Medzihorsky, J., Pernes, J., Von Römer, J., Stepanova, N., Sundström, A., Tzelgov, E., Wang, J-T., Wig, T., and Ziblatt, D. (2019B). «V-Dem Codebook v9» Varieties of Democracy (V-Dem) Project.

Cunningham, D. E., Gleditsch, K.S. & Saleyhan, I. (2009). It takes two: A dyadic analysis of civil war duration and outcome. *Journal of Conflict Resolution*, 53(4):570–597.

Cunningham, K. & Weidmann, N. (2010). Shared space: Ethnic groups, state accommodation and localized conflict. *International Studies Quarterly*, 54:1035-1054.

Denny, E. & Walter, B. (2014). Ethnicity and civil war. *Journal of peace research*, 52(2):199-212.

De Soysa, I. (2002). Ecoviolence: shrinking pie or honey pot? *Global Environmental Politics*, 2(4):1-36.

Easterly, W. & Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The quarterly journal of economics* 112(3):1203-1250.

Elbadawi, I. & Sambanis, N. (2002) How much war will we see?: Explaining the prevalence of civil war. *Journal of conflict resolution*, 46(3):307-334

Esteban, J. & Ray, D. (1999). Conflict and distribution. *Journal of economic theory*, 87(2):379-415.

European Food Safety Authority (2010). Scientific opinion on dietary reference values for water. *EFSA journal*, 8(3):1459.

Ellingsen, T. (2000). Colorful community or ethnic witches brew? *Journal of conflict resolution*, 44(2):228-249.

Fearon, J. & Laitin, D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, 97(1):75-90.

Fjelde, H. (2009). Buying peace? Oil wealth, corruption and civil war, 1985-99. *Journal of peace research*, 46(2):199-218.

Fjelde, H. & Østby, G. (2014). Socioeconomic inequality and communal conflict: A disaggregated analysis of sub-saharan africa, 1990-2008. *International interactions*, 40(5):737-762.

Fluckiger, M. & Ludvig, M. (2018). Youth bulges and civil conflict: causal evidence from Sub-Saharan Africa. *Journal of conflict resolution*, 62(9):1932-1962.

FN-Sambandet (2020). «Befolkningstall» lastet ned 12.03.20 fra <https://www.fn.no/Statistikk/Befolkningstall>

Gates, S. (2002). Recruitment and Allegiance: The Microfoundations of Rebellion, *Journal of Conflict Resolution* 46(1):111–130.

Gizelis, T-I. & Wooden, A.E. (2010) Water Resources, Institutions, and Intrastate Conflict, *Political Geography* 29(8):444–453.

Gleditsch, N.P., Wallensteen, P., Eriksson, M., Sollenberg, M. & Strand, H. (2002). Armed Conflict 1946–2001: A New Dataset. *Journal of Peace Research*, 39(5):615–637.

Gleick, P. (1993). Water and conflict: Fresh water resources and international security. *International security*, 18(1):79-112.

Gurr, T-R. (1993). Why minorities rebel: A global analysis of communal mobilization and conflict since 1945. *International Political science review*, 14(2):161-201.

Gurr, T-R (1994). Peoples against states: Ethnopolitical conflict and the changing world system: 1994 presidential address. *International studies quarterly*, 38(3):347-377.

Hauge, W. & Ellingsen, T. (2001) Causal pathways to peace. I P.F. Diehl, N.P. Gleditsch (Eds.), *Environmental conflict*, Westview Press:26-57.

Hegre, H., Ellingsen, T., Gates, S & Gleditsch, N.P. (2001). Toward a democratic civilpeace? Democracy, political change, and civil war, 1816-1992. *American Political Science Review*, 95:33-48.

Hegre, H., & Sambanis, N. (2006). Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution*, 50(4):508–535.

Herbst, J. (2002). States and power in Africa: Comparative lessons in authority and control. Princeton, NJ: Princeton University Press.

Homer-Dixon, T. (1999). Environment, Scarcity, and Violence. Princeton, NJ & Oxford: Princeton University Press.

- Horowitz, D. (1985). *Ethnic groups in conflict*. Berkeley: Univ Calif. Press.
- Humphreys, M. & Weinstein, J. (2008). Who fights? The determinants of participation in civil war. *American Journal of Political Science*, 52(2):436–455.
- Lake, D.A. & Rotchild, D. (1998). *The international spread of ethnic conflict*. Princeton: Princeton University Press.
- Le Billon, P. (2003). Buying peace or fuelling war: the role of corruption in armed conflicts. *J. Int. Dev.*, 15 413-426.
- Linke, A., Witmer, F., Holland, E. & O'Loughlin, J. (2017). Mountainous Terrain and Civil Wars: Geospatial Analysis of Conflict Dynamics in the Post-Soviet Caucasus, *Annals of the American Association of Geographers*, 107(2):520-535.
- Manuelli, R.E. & Seshadri, A. (2014). Human capital and the wealth of nations. *Am Econ Rev.*, 103(9):2736-2762.
- McFerson, H.M. (2009). Governance and Hyper-corruption in Resource-rich African Countries. *Third World Quarterly*, 30(8):1529-1547.
- Mehmetoglu, M. & Jakobsen, T.G. (2017). *Applied statistics using stata*. London: Sage publications.
- Midtbø, T. (2016). *Stata. En entusiastisk innføring*. Oslo: Universitetsforlaget.
- Olson, M. (1965). *The theory of collective action: public goods and the theory of groups*. Cambridge: *Harvard University Press*.
- Pettersson, T. (2019). UCDP/PRIO Armed Conflict Dataset Codebook v 19.1 (<https://ucdp.uu.se/downloads/>).
- Pickering, S. (2012). Proximity, maps and conflict: New measures, new maps and new findings. *Conflict Management and Peace Science*, 29(4):425–430.

- Raleigh, C. & Hegre, H. (2009). Population size, concentration, and civil war: a geographically disaggregated analysis. *Political geography*, 28(4):224-238.
- Reynal-Querol, M. (2002). Ethnicity, Political Systems, and Civil War. *Journal of Conflict Resolution*, 46(1):29-54.
- Ringdal, K. (2013). *Enhet og Mangfold*. Bergen: Fagbokforlaget.
- Sambanis, N. (2004). What Is Civil War?: Conceptual and Empirical Complexities of an Operational Definition. *Journal of Conflict Resolution*, 48(6):814–858.
- Scarcelli, M. (2017). The Uneven Application of the ‘Civil War’ Label to Iraq, *Civil Wars*, 19(1):87-107.
- Skog, O. (1998). *Å forklare sosiale fenomener: en regresjonsbasert tilnærming*. Oslo: Ad notam Gyldendal.
- Skog, O. (2017). *Å forklare sosiale fenomener: en regresjonsbasert tilnærming*. Oslo: Ad Notam Gyldendal.
- Smith, M.J. (1993). *Power, pressure and policy: state autonomy and policy networks in Britain and the United states*. Pittsburgh, PA: University of Pittsburgh Press.
- Sommers, M. (2003). *War, urbanization and Africas youth at risk: understanding and addressing future challenges*. Washington, DC: Basic education and policy support activity and creative associates international.
- Taylor, M. (1988). “Rationality and Revolutionary Collective Action.”. In *Rationality and Revolution*, ed. Michael Taylor. New York: Cambridge University Press, 63–91.
- Tir, J. & Diehl, P.F. (2001). Demographic pressure and interstate conflict: linking population growth and density to Militarized disputes and wars. I P.F. Diehl, N.P. Gleditsch (Eds.), *Environmental conflict*, Westview Press, Boulder, 58-83.

Theisen, O.M., Strand, H. & Østby, G. (2020) Ethno-political favouritism in maternal health care service delivery: Micro-level evidence from sub-Saharan Africa, 1981-2014.

International area studies review, 23(1):3-27.

Theobald, R. (1990). *Corruption, Development and Underdevelopment*. Macmillan: Basingstoke.

Toft, M. (2014). Territory and war. *Journal of peace research*, 51(2):185-198

Tucker, R. (1999). BTSCS: «A Binary Time-Series-Cross-Section Data Analysis Utility». Version 4.0.4 Cambridge, MA: Harvard University.

Uppsala Conflict Data program (udatert). Countries in conflict view. Hentet 15.2.2020 fra <https://ucdp.uu.se/encyclopedia>

Urdal, H. (2011). Demography and armed conflict: assessing the role of population growth and youth bulges. Centre for research on peace and development.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P., & Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family. *Journal of Conflict Resolution*, 59(7):1327–1342.

World Bank (2020). World Bank Indicators. Hentet 3.2.2020 fra <https://data.worldbank.org/>

Yair, O. & Miodownik, D. (2016). Youth bulge and civil war: why a countrys share of young adults explains only non-ethnic wars. *Conflict management and peace science*, 33(1):25-44

7. Appendiks

7.1 Liste over land-år kodet som intern konflikt i Onset-datasettet, men ikke i Armed Conflict-datasettet.

Algeria 1963	Cambodia 2011	China 1978	Egypt 1973	Indonesia 1963	Portugal 1964
Argentina 1982	Cameroon 1996	Cyprus 1974	El Salvador 1969	Iran 1974	Tanzania 1978
Australia 2003	China 1962	Djibouti 2008	Eritrea 1998	Laos 1986	United Kingdom 1962
Burkina Faso 1985	China 1969	Ecuador 1995	Eritrea 2016	Portugal 1961	United Kingdom 1964
Cambodia 1977	China 1974	Egypt 1967	Indonesia 1962	Portugal 1963	United Kingdom 2003

7.2 Test for multikollinearitet for modellene 1-8 (VIF)

Modell 1

Variabel	VIF	1/VIF
Fredsår	1,25	0,803158
(ln) bnp p.capita	1,22	0,818648
(ln) befolkning	1,02	0,978235
Mean VIF	1,16	

Modell 2

Variabel	VIF	1/VIF
(ln) areal	1,90	0,526147
(ln) befolkning	1,87	0,535323
fredsår	1,26	0,794300
(ln) bnp p.capita	1,22	0,817659
Mean VIF	1,56	

Modell 3

Variabel	VIF	1/VIF
Relevante grupper	1,31	0,765950
(ln) befolkning	1,27	0,785330
fredsår	1,25	0,800601
(ln) bnp p.capita	1,25	0,802539
Mean VIF	1,27	

Modell 4

Variabel	VIF	1/VIF
fredsår	1,25	0,802359
(ln) befolkning	1,23	0,815386
(ln) bnp p.capita	1,23	0,815806
Ekskluderte grupper	1,22	0,820074
Mean VIF	1,23	

Modell 5

Variabel	VIF	1/VIF
(ln) befolkning	82,72	0,012088
(ln) unge menn	82,37	0,012140
(ln) bnp p.capita	1,26	0,792286
fredsår	1,25	0,799528
Mean VIF	41,90	

Modell 6

Variabel	VIF	1/VIF
(ln) rural befolkning	2,81	0,356273
(ln) befolkning	2,51	0,397982
(ln) bnp p.capita	1,43	0,701008
fredsår	1,25	0,800389
Mean VIF	2,00	

Modell 7

Variabel	VIF	1/VIF
(ln) bnp p.capita	1,86	0,537632
Korrupsjon	1,73	0,579443
fredsår	1,25	0,798026
(ln) befolkning	1,03	0,970323
Mean VIF	1,47	

Modell 8

Variabel	VIF	1/VIF
(ln) areal	14,74	0,067849
(ln) befolkning	9,03	0,110704
fredsår	6,71	0,149034
(ln) bnp p.capita	1,55	0,646917
Mean VIF	8,01	

7.3 Betydningsfulle enheter med cooks D for modell 1.

Afghanistan 2015 0,0031	Angola 2017 0,0028	Azerbaijan 2012 0,0019	Burundi 2014 0,0019	Chad 1983 0,0022	Comoros 1997 0,0086	Congo dem rep 2013 0,0048	Cote d'Ivoire 2011 0,0024	Eritrea 1997 0,0034
Algeria 1991 0,0027	Argentina 1963 0,0024	Bangladesh 1975 0,0047	Cameroon 1984 0,0019	Chad 1987 0,0021	Congo dem.rep 1961 0,0021	Congo dem rep 2017 0,0051	Dominican Republic 1965 0,0024	Eritrea 2003 0,0031
Algeria 2015 0,0034	Argentina 1974 0,0026	Bangladesh 2005 0,0068	Cameroon 2015 0,0025	Chad 1997 0,0019	Congo dem.rep 1964 0,0021	Congo rep 1993 0,0031	Egypt 1993 0,0043	Ethiopia 1983 0,0033
Angola 1991 0,0197	Azerbaijan 1991 0,0019	Bangladesh 2016 0,0073	Central African Republic 2001 0,0025	Chad 2015 0,0020	Congo dem.rep 1977 0,0025	Congo rep 1997 0,0029	Egypt 2014 0,0054	Ethiopia 1991 0,0038
Angola 1994 0,0020	Azerbaijan 1993 0,0019	Bolivia 1967 0,0023	Central African Republic 2006 0,0023	Chile 1973 0,0019	Congo dem.rep 1996 0,0035	Congo rep 2002 0,0026	Egypt 2015 0,0055	Ethiopia 1993 0,0040
Angola 1998 0,0021	Azerbaijan 1997 0,0019	Burkina Faso 1987 0,0019	Central African Republic 2015 0,0025	China 2008 0,0173	Congo dem.rep 2006 0,0042	Congo Rep 2016 0,0021	El Salvador 1972 0,0024	Ethiopia 1996 0,0043
Angola 2002 0,0022	Azerbaijan 2005 0,0019	Burundi 1965 0,0027	Chad 1966 0,0026	Colombia 1964 0,0022	Congo dem rep. 2007 0,0042	Cote d'Ivoire 2002 0,0021	El Salvador 1979 0,0022	Ethiopia 1998 0,0044
Angola 2007 0,0023	Azerbaijan 2008 0,0019	Burundi 1991 0,0020	Chad 1976 0,0023	Comoros 1989 0,0098	Congo dem rep 2011 0,0046	France 1961 0,0037	Gabon 1964 0,0084	Gambia 1981 0,0074

Georgia 1991 0,0022	Guinea - Bissau 1998 0,0052	India 1982 0,014	India 1992 0,015	India 2009 0,016	Indonesia 1981 0,0071	Iran 1980 0,0033	Irak 1973 0,0019	Israel 2000 0,0020	Lesotho 1998 0,0036
Georgia 1992 0,0022	Haiti 1989 0,0019	India 1983 0,014	India 1993 0,015	India 2012 0,017	Indonesia 1984 0,0074	Iran 1986 0,0038	Irak 1982 0,0020	Israel 2006 0,0019	Liberia 2000 0,0029

Georgia 2004 0,0024	Haiti 2004 0,0019	India 1984 0,014	India 1994 0,015	India 2013 0,017	Indonesia 1990 0,0078	Iran 1993 0,0043	Irak 1987 0,0021	Jordan 2016 0,0019	Libya 2011 0,0020
Georggia 2008 0,0024	India 1964 0,012	India 1987 0,014	India 1996 0,015	India 2013 0,017	Indonesia 1992 0,0079	Iran 1996 0,0044	Irak 1990 0,0022	Kenya 1982 0,0022	Libya 2014 0,0020
Ghana 1966 0,0019	India 1966 0,012	India 1989 0,014	India 1997 0,015	India 2014 0,017	Indonesia 1997 0,0082	Iran 1997 0,0044	Irak 2004 0,0027	Kenya 2015 0,0038	Libya 2015 0,0019
Ghana 1091 0,0019	India 1969 0,012	India 1990 0,015	India 2000 0,016	India 2015 0,017	Indonesia 1999 0,0084	Iran 2005 0,0047	Israel 1967 0,0030	Laos 1989 0,0023	Malaysi a 1963 0,0019
Guatemala 1963 0,0022	India 1971 0,012	India 1991 0,015	India 2003 0,016	Indonesia 1975 0,0066	Iran 1966 0,0026	Iran 2016 0,0051	Israel 1973 0,0027	Lebano n 1989 0,0030	Malaysi a 1974 0,0019
Guinea 2000 0,0019	India 1979 0,013	India 1991 0,015	India 2008 0,016	Indonesia 1076 0,0067	Iran 1979 0,0032	Iran 2017 0,0051	Israel 1990 0,0022	Lebano n 2014 0,0020	Malaysi a 1981 0,0020

Malay sia 1974 0,0019	Mali 2015 0,0022	Mozambi que 2016 0,0027	Nepal 1996 0,0024	Niger 2015 0,0023	Pakist an 1971 0,0043	P-N Guinea 1990 0,0022	Russi a 1990 0,0071	Rwan da 2016 0,0019	Spania 1985 0,0033
Malay sia 1981 0,0020	Mali 2017 0,0022	Myanmar 2000 0,0037	Nicarag ua 1977 0,0028	Nigeria 1966 0,0039	Pakist an 1974 0,0045	Paragua y 1989 0,0023	Russi a 1993 0,0071	Saudi - Arabi a 1979 0,0019	Spania 1991 0,0033
Malay sia 2013 0,0020	Mauritani a 1975 0,0047	Myanmar 2005 0,0038	Nicarag ua 1982 0,0028	Nigeria 1967 0,0040	Pakist an 1990 0,0060	Peru 1965 0,0019	Russi a 1994 0,0071	Seneg al 1990 0,0019	Sri Lanka 1971 0,0019
Mali 1990 0,0019	Mauritani a 2010 0,0026	Myanmar 2009 0,0039	Niger 1991 0,0019	Nigeria 2004 0,0067	Pakist an 1994 0,0063	Peru 1982 0,0022	Russi a 1999 0,0070	Seneg al 2011 0,0019	Sri Lanka 1984 0,0021
Mali 1994	Mexico 1994 0,0054	Myanmar 2011 0,0039	Niger 1994 0,0019	Nigeria 2009 0,0072	Pakist an 2004	Peru 2007 0,0028	Russi a 2007	Sierra Leone 1991	Sri Lanka

0,0019					0,0073		0,0069	0,0023	1989 0,0021
Mali 2007 0,0020	Morocco 1971 0,0021	Myanmar 2013 0,0040	Niger 1995 0,0019	Nigeria 2015 0,0078	Pakistan 2007 0,0075	Philippines 1969 0,0031	Russia 2015 0,0070	Sør Afrika 1966 0,0023	Suriname 1987 0,010
Mali 2009 0,0020	Morocco 1075 0,0022	Myanmar 2014 0,0040	Niger 1997 0,0019	North Macedonia 2001 0,0036	Pakistan 2016 0,0083	Philippines 1970 0,0032	Rwanda 1990 0,0019	Sør Afrika 1981 0,0028	Tajikistan 1992 0,0029
Mali 2012 0,0021	Mozambique 2013 0,0026	Myanmar 2016 0,0040	Niger 2007 0,0020	Oman 1069 0,0071	Panama 1989 0,0032	Philippines 1993 0,0046	Rwanda 2009 0,0019	Spania 1978 0,0032	Tajikistan 2010 0,0019

Thailand 1974 0,0034	Tunisia 2016 0,0019	Turkey 2016 0,0051	United States 2003 0,0097	Venezuela 1982 0,0021	Zimbabwe 1967 0,0022
Thailand 2003 0,0045	Turkey 1984 0,0038	Ukraine 2014 0,0036	Uruguay 1972 0,0029	Venezuela 1992 0,0023	Zimbabwe 1973 0,0020
Togo 1986 0,0026	Turkey 1991 0,0041	United Kingdom 1971 0,0041	Uzbekistan 1999 0,0025	Yemen 1994 0,0020	
Trinidad og Tobago 1990 0,0050	Turkey 2005 0,0046	United Kingdom 1998 0,0042	Uzbekistan 2004 0,0026	Yemen 2009 0,0024	
Tunisia 1980 0,0020	Turkey 2015 0,0050	United States 2001 0,0096	Venezuela 1962 0,0019	Yemen 2015 0,0027	

7.4 Liste over uteliggere med standardiserte residualer større enn 3 for modell 1.

Land standardisert residual år

Afghanistan	5.101204	2015
Algeria	5.123	1991
Algeria	5.089404	2015
Angola	5.186251	1991
Angola	5.178209	1994
Angola	5.167595	1998
Angola	5.156755	2002
Angola	5.142205	2007
Angola	5.112991	2017
Argentina	5.139938	1963
Argentina	5.125976	1974
Azerbaijan	5.229301	1991
Azerbaijan	5.226794	1993
Azerbaijan	5.223095	1997

Azerbaijan	5.217459	2005
Azerbaijan	5.213882	2008
Azerbaijan	5.209013	2012
Bangladesh	5.042834	1975
Bangladesh	4.986695	2005
Bangladesh	4.976249	2016
Bolivia	5.274514	1967
Burkina Faso	5.219975	1987
Burundi	5.299986	1965
Burundi	5.251404	1991
Burundi	5.204281	2014
Cameroon	5.205009	1984
Cameroon	5.133284	2015
C.A.R	5.284688	2001
C.A.R	5.276316	2006
C.A.R	5.272016	2009
Chad	5.292863	1966
Chad	5.275216	1976
Chad	5.263079	1983
Chad	5.253695	1987
Chad	5.226701	1997
Chad	5.174583	2015
Chile	5.200821	1973
China	4.802825	2008
Colombia	5.153728	1964
Comoros	5.470233	1989
Comoros	5.451207	1997
Congo, D.R	5.166116	1961
Congo, D.R.	5.159702	1964
Congo, D.R.	5.129993	1977
Congo, D.R.	5.083368	1996
Congo, D.R.	5.060374	2006
Congo, D.R.	5.057698	2007
Congo, D.R	5.046804	2011
Congo, D.R.	5.041335	2013
Congo, D.R.	5.030542	2017
Congo, Rep.	5.315705	1993
Congo, Rep.	5.306303	1997
Congo, Rep.	5.294395	2002
Congo, Rep.	5.260572	2016
Cote d'Ivoire	5.15812	2002
Cote d'Ivoire	5.141721	2011
Dom. Rep.	5.281293	1965
Egypt	5.055719	1993
Egypt	5.021924	2014
Egypt	5.020116	2015
Egypt	5.016645	2017
El Salvador	5.281517	1972
El Salvador	5.26877	1979
Eritrea	5.328453	1997

Eritrea	5.314384	2003
Ethiopia	5.09275	1983
Ethiopia	5.071162	1991
Ethiopia	5.065279	1993
Ethiopia	5.057097	1996
Ethiopia	5.052197	1998
France	5.07518	1961
Gabon	5.447479	1964
Gambia, The	5.428689	1981
Georgia	5.263018	1991
Georgia	5.262377	1992
Georgia	5.280241	2004
Georgia	5.281922	2008
Ghana	5.222015	1966
Ghana	5.19255	1981
Guatemala	5.267237	1963
Guinea	5.21896	2000
Guinea-Bissau	5.381839	1998
Haiti	5.233655	1989
Haiti	5.211252	2004
India	4.883993	1964
India	4.880599	1966
India	4.875416	1969
India	4.871778	1971
India	4.856737	1979
India	4.851067	1982
India	4.849177	1983
India	4.847307	1984
India	4.84188	1987
India	4.838414	1989
India	4.836724	1990
India	4.835064	1991
India	4.833433	1992
India	4.831829	1993
India	4.830248	1994
India	4.827144	1996
India	4.825623	1997
India	4.821218	2000
India	4.817093	2003
India	4.810818	2008
India	4.80967	2009
India	4.806522	2012
India	4.805561	2013
India	4.804629	2014
India	4.80372	2015
Indonesia	4.991767	1975
Indonesia	4.989735	1976
Indonesia	4.979974	1981
Indonesia	4.97443	1984
Indonesia	4.964932	1990

Indonesia	4.962129	1992
Indonesia	4.955811	1997
Indonesia	4.95351	1999
Iran	5.125454	1966
Iran	5.094724	1979
Iran	5.091664	1980
Iran	5.071743	1986
Iran	5.055934	1993
Iran	5.052476	1996
Iran	5.051376	1997
Iran	5.04319	2005
Iran	5.032412	2016
Iran	5.031276	2017
Iraq	5.195491	1973
Iraq	5.172726	1982
Iraq	5.162916	1987
Iraq	5.157228	1990
Iraq	5.12327	2004
Israel	5.309916	1967
Israel	5.29521	1973
Israel	5.266084	1990
Israel	5.241291	2000
Israel	5.231808	2006
Jordan	5.206774	2016
Kenya	5.155742	1982
Kenya	5.074079	2015
Lao PDR	5.275899	1989
Lebanon	5.310478	1989
Lebanon	5.241641	2014
Lesotho	5.336283	1998
Liberia	5.306849	2000
Libya	5.241839	2011
Libya	5.240336	2014
Libya	5.239609	2015
Malaysia	5.211924	1963
Malaysia	5.188835	1974
Malaysia	5.174448	1981
Malaysia	5.113956	2013
Mali	5.21689	1990
Mali	5.208667	1994
Mali	5.177312	2007
Mali	5.17188	2009
Mali	5.164335	2012
Mali	5.157135	2015
Mali	5.152215	2017
Mauritania	5.370028	1975
Mauritania	5.289919	2010
Mexico	5.022168	1994
Morocco	5.162442	1971
Morocco	5.155464	1975

Mozambique	5.125659	2013
Mozambique	5.118662	2016
Myanmar	5.07609	2000
Myanmar	5.072261	2005
Myanmar	5.070109	2009
Myanmar	5.068908	2011
Myanmar	5.067533	2013
Myanmar	5.066857	2014
Myanmar	5.065666	2016
Nepal	5.137666	1996
Nicaragua	5.302955	1977
Nicaragua	5.29085	1982
Niger	5.21848	1991
Niger	5.210153	1994
Niger	5.207304	1995
Niger	5.201529	1997
Niger	5.171294	2007
Niger	5.145843	2015
Nigeria	5.068543	1966
Nigeria	5.066761	1967
Nigeria	4.988912	2004
Nigeria	4.978159	2009
Nigeria	4.965056	2015
Macedonia	5.334421	2001
Oman	5.423393	1969
Pakistan	5.05592	1971
Pakistan	5.049108	1974
Pakistan	5.007642	1990
Pakistan	4.9985	1994
Pakistan	4.976928	2004
Pakistan	4.971307	2007
Pakistan	4.955488	2016
Panama	5.320385	1989
P.N. Guinea	5.266871	1990
Paraguay	5.27643	1989
Peru	5.189952	1965
Peru	5.152598	1982
Peru	5.117189	2007
Philippines	5.100319	1969
Philippines	5.09795	1970
Philippines	5.047002	1993
Philippines	5.010731	2016
Russia	4.981421	1990
Russia	4.981296	1993
Russia	4.981397	1994
Russia	4.982018	1999
Russia	4.984507	2007
Russia	4.983769	2015
Rwanda	5.229098	1990
Rwanda	5.204799	2009

Rwanda	5.190253	2016
Saudi Arabia	5.210427	1979
Senegal	5.22645	1990
Senegal	5.181129	2011
Sierra Leone	5.271806	1991
South Africa	5.146311	1966
South Africa	5.114336	1981
Spain	5.09555	1978
Spain	5.092049	1985
Spain	5.090994	1991
Sri Lanka	5.182959	1971
Sri Lanka	5.164495	1984
Sri Lanka	5.158719	1989
Suriname	5.474212	1987
Tajikistan	5.252337	1992
Tajikistan	5.226438	2010
Thailand	5.086311	1974
Thailand	5.04956	2003
Togo	5.293088	1986
Trinidad/Tobago	5.377165	1990
Tunisia	5.24018	1980
Tunisia	5.192873	2016
Turkey	5.073675	1984
Turkey	5.062934	1991
Turkey	5.045405	2005
Turkey	5.033484	2015
Turkey	5.032146	2016
Ukraine	5.078675	2014
U.K	5.061369	1971
U.K	5.057651	1998
United States	4.92803	2001
United States	4.92657	2003
Uruguay	5.307638	1972
Uzbekistan	5.129781	1999
Uzbekistan	5.124687	2004
Venezuela	5.214077	1962
Venezuela	5.164139	1982
Venezuela	5.143585	1992
Yemen, Rep.	5.1735	1994
Yemen, Rep.	5.136094	2009
Yemen, Rep.	5.122697	2015
Zimbabwe	5.263983	1967
Zimbabwe	5.246882	1973

7.5 Modeller uten de to mest betydningsfulle enhetene

Tabell 9 viser modell 18-25, som alle er kjørt på datasettet uten de betydningsfulle enhetene «Kina 2008» og «India 2013». Som man kan lese av tabellen er koeffisientene på samtlige variabler endret, men det er såpass lite at det ikke vil bli kommentert her. De største endringene er imidlertid at modell 23 som inkluderer rural befolkning har fått en befolkningskoeffisient som er høyere enn den er i modell 18, som vil si at størrelsen på den rurale befolkningen i landet ikke lenger stjeler forklaringskraft fra befolkningsvariabelen. I modell 25 har befolkningsvariabelen mistet signifikansnivået sitt, som gjør at resultatet er mindre til å stole på enn hva det var i Tabell 5. Bortsett fra dette er det ingen store endringer i modellene sammenliknet med Tabell 5.

	(18)	(19)	(20)	(20b)	(21)	(21b)	(22)	(23)	(24)	(25)	(25b)
	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2
(ln) befolkning	0.351** *	0.262***	0.281***	0.229***	0.340***	0.294***	0.351***	0.352***	0.362***	0.225	0.283*
	(0.040 0)	(0.0540)	(0.0501)	(0.0519)	(0.0470)	(0.0496)	(0.0409)	(0.104)	(0.0413)	(0.117)	(0.119)
(ln) areal		0.132* (0.0549)									
Relevante grupper			0.0240* (0.0104)	0.129*** (0.0323)							
Relevante grupper^2				- 0.00278** *							
				(0.00084 2)							
Ekskluderte grupper					0.00527 (0.0113)	0.103** (0.0395)					
Ekskluderte grupper^2						-0.00268* (0.00106)					
(ln) unge menn							0.0378 (0.641)				
(ln) rural befolkning								-0.00123 (0.0934)			
Korrupsjon									0.780** (0.294)		0.747* (0.298)
(ln) BNP										0.120 (0.105)	0.0745 (0.105)
fredsår	0.0502 (0.067 5)	0.0443 (0.0677)	0.0453 (0.0677)	0.0581 (0.0682)	0.0483 (0.0677)	0.0584 (0.0680)	0.0505 (0.0677)	0.0503 (0.0675)	0.0633 (0.0676)	0.0657 (0.0687)	0.0721 (0.0686)
(ln) BNP p.cap	- 0.253** *	- 0.266***	-0.239***	-0.203***	-0.252***	-0.237***	-0.254***	-0.253***	-0.192***	-0.363***	-0.263* (0.116)
	(0.050 0)	(0.0505)	(0.0503)	(0.0516)	(0.0500)	(0.0503)	(0.0516)	(0.0588)	(0.0566)	(0.109)	
_cons	- 6.681** *	- 6.784***	-5.732***	-5.590***	-6.512***	-6.106***	-6.748***	-6.679***	-7.826***	-6.571***	-7.707*** (0.938)
	(0.798)	(0.799)	(0.900)	(0.899)	(0.877)	(0.885)	(1.393)	(0.809)	(0.923)	(0.805)	
<i>N</i>	7121	7121	7121	7121	7121	7121	7121	7121	7121	7121	7121
pseudo <i>R</i> ²	0.102	0.105	0.104	0.110	0.102	0.105	0.102	0.102	0.105	0.103	0.106

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Modellen inneholder også tre splines som ikke rapporteres

Tabell 9: Modell 18-25 uten Kina 2008 og India 2013.

7.6 Modeller kjørt på datasettet delt i to tidsserier (1961-1990 og 1991-2017)

Tabell 10 og 11 inneholder som sagt regresjonsanalyser der datasettet er delt til før og etter

den kalde krigen. Tabell 10 sitt datagrunnlag er 29 år med 2973 observasjoner, mens Tabell

11 sitt datagrunnlag er 26 år med 4150 observasjoner. Ettersom verden ser ganske annerledes

ut i dag kontra hva den har gjort tidligere, vil et slikt skille kunne gi oss et interessant innblikk i hvordan sammenhengen mellom befolkningsstørrelse og konflikt kan forklares ytterligere.

Til sammenlikning med Tabell 5, som er modellene kjørt på hele datasettet, er ikke befolkningsvariabelen signifikant i modell 30, 31, 33 eller 33b i Tabell 10. Modell 33 inneholder BNP som i Tabell 5 stjal mye forklaringskraft fra befolkning. Denne var ikke signifikant da Kina og India ble tatt bort, og er heller ikke signifikant i denne modellen. Modell 31, der rural befolkning er inkludert, var tidligere signifikant i Tabell 5, men er ikke det i Tabell 10. For modell 28b og 29b, som inkluderer etniske grupper, viser det seg også å være en endring i det kurvelineære forholdet som man kunne se i Tabell 6. For årene 1961-1990 er det ikke et signifikant kurvelineært forhold mellom verken befolkning og relevante grupper eller befolkning og ekskluderte grupper. Som man kan se er heller ikke antall ekskluderte grupper signifikant lenger, verken med eller uten andregradsleddet. Inkluderingen av ekskluderte grupper gjør også at befolkningsvariabelens koeffisient øker sammenliknet med modell 26.

Modellene 28 og 28b, som tester påvirkningen antall relevante etniske grupper har på sammenhengen mellom befolkning og konflikt, har heller ingen signifikante verdier på variabelen som måler antallet etniske grupper. Befolkningsvariabelen er, om ikke annet, signifikant på 0,05-nivå, og har sunket i begge modellene sammenliknet med modell 26. Modell 28 har henholdsvis en befolkningskoeffisient som har sunket med 27,9% sammenliknet med modell 15, mens modell 28b har en nedgang på 40,1% i befolkningskoeffisienten. Dette gjør at inkluderingen av relevante etniske grupper har en stor påvirkning på sammenhengen mellom konflikt og befolkningsstørrelse for perioden 1961-1990 sammenliknet med påvirkningen vi har sett i de andre tabellene, men variabelen som måler etniske grupper er ikke signifikant.

	(26)	(27)	(28)	(28b)	(29)	(29b)	(30)	(31)	(32)	(33)	(33b)
	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2
(ln) befolkning	0.269***	0.249**	0.194*	0.161*	0.289***	0.281***	0.295	0.193	0.281***	0.00503	0.00340
	(0.0598)	(0.0812)	(0.0756)	(0.0786)	(0.0730)	(0.0781)	(0.879)	(0.190)	(0.0618)	(0.209)	(0.207)
(ln) areal		0.0299									
		(0.0805)									
Relevante grupper			0.0296	0.101*							
			(0.0184)	(0.0499)							
Relevante grupper^2				-0.00191							
				(0.00135)							
Ekskluderte grupper					-0.0107	0.00536					
					(0.0230)	(0.0646)					
Ekskluderte grupper^2							-0.000455				
							(0.00174)				
(ln) unge menn							-0.0254				
							(0.863)				
(ln) rural befolkning								0.0764			
								(0.181)			
Korrupsjon									0.489		0.534
									(0.419)		(0.422)
(ln) BNP										0.257	0.272
										(0.195)	(0.193)
fredsår	-0.143	-0.147	-0.152	-0.145	-0.139	-0.138	-0.143	-0.144	-0.129	-0.0963	-0.0774
	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.121)	(0.121)
(ln) BNP p.cap	-0.221**	-0.223**	-0.180*	-0.153	-0.231**	-0.228**	-0.221**	-0.197*	-0.179*	-0.438*	-0.403*
	(0.0740)	(0.0743)	(0.0785)	(0.0809)	(0.0770)	(0.0779)	(0.0756)	(0.0922)	(0.0841)	(0.181)	(0.182)
_cons	-5.635***	-5.663***	-4.817***	-4.776***	-5.868***	-5.805***	-5.729	-5.766***	-6.435***	-5.545***	-6.424***
	(1.221)	(1.223)	(1.323)	(1.322)	(1.317)	(1.335)	(3.431)	(1.262)	(1.420)	(1.225)	(1.431)
<i>N</i>	2973	2973	2973	2973	2973	2973	2973	2973	2973	2973	2973
pseudo <i>R</i> ²	0.058	0.059	0.061	0.064	0.059	0.059	0.058	0.059	0.060	0.060	0.062

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Modellen inneholder også 3_splines som ikke rapporteres

Tabell 10: Modell 26-33, 1961-1990

For modellene kjørt på årstallene 1991-2017 har det også skjedd et par endringer sammenliknet med Tabell 5. Likevel er det flere resultater som også viser i samme retning. Modell 35, 36, 36b, 37 og 37b viser signifikante verdier for befolkningsvariabelen, og er lavere enn befolkningsvariabelen i modell 34. Areal og andregradsleddene for etniske grupper og ekskluderte grupper er også signifikante.

Både modell 39, 40, 41 og 41b har signifikant befolkningsvariabel, men til forskjell fra Tabell 5 der det bare er modellen med korrupsjon som viste en høyere koeffisient sammenliknet med modell 1, viser alle fire modellene høyere koeffisient på befolkningsvariabelen sammenliknet med modell 34. En høyere koeffisient taler imot at den inkluderte variabelen kan forklare sammenhengen mellom befolkning og konflikt. Inkluderingen av størrelse på rural befolkning, korrupsjonsnivå og BNP gjør i disse modellene at befolkningsstørrelse har en sterkere påvirkning på konflikt enn tidligere.

	(34)	(35)	(36)	(36b)	(37)	(37b)	(38)	(39)	(40)	(41)	(41b)
	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2	onset2
(ln) befolkning	0.434*** (0.0554)	0.268*** (0.0737)	0.359*** (0.0678)	0.279*** (0.0706)	0.405*** (0.0628)	0.330*** (0.0654)	0.507 (0.716)	0.503*** (0.139)	0.437*** (0.0561)	0.759*** (0.230)	0.732** (0.227)
(ln)areal		0.251** (0.0765)									
Relevante grupper			0.0232 (0.0125)	0.163*** (0.0422)							
Relevante grupper^2				- 0.00349*** (0.00103)							
Ekskluderte grupper					0.0120 (0.0125)	0.174*** (0.0481)					
Ekskluderte grupper^2						- 0.00422*** (0.00122)					
(ln)unge menn							-0.0722 (0.697)				
(ln) rural befolkning								-0.0661 (0.120)			
korruptsjon									0.830 (0.475)		0.791 (0.478)
(ln) BNP										-0.316 (0.216)	-0.287 (0.214)
fredsår	0.153 (0.0891)	0.169 (0.0894)	0.150 (0.0893)	0.179* (0.0907)	0.148 (0.0892)	0.175 (0.0904)	0.153 (0.0891)	0.154 (0.0891)	0.150 (0.0891)	0.152 (0.0893)	0.148 (0.0894)
(ln) bnp p.cap	- 0.308*** (0.0693)	- 0.341*** (0.0705)	- 0.315*** (0.0699)	-0.274*** (0.0709)	- 0.316*** (0.0699)	-0.304*** (0.0699)	- 0.307*** (0.0710)	- 0.332*** (0.0821)	-0.243** (0.0798)	0.00539 (0.225)	0.0392 (0.224)
_cons	- 7.634*** (1.047)	- 7.870*** (1.057)	- 6.479*** (1.222)	-6.053*** (1.228)	- 7.137*** (1.172)	-6.407*** (1.190)	-7.910** (2.865)	- 7.566*** (1.055)	- 8.756*** (1.247)	- 7.901*** (1.065)	- 8.937*** (1.252)
N	4150	4150	4150	4150	4150	4150	4150	4150	4150	4150	4150
pseudo R ²	0.148	0.157	0.151	0.160	0.149	0.158	0.148	0.149	0.151	0.150	0.152

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Modellen inneholder også 3_splines som ikke rapporteres

Tabell 11: Modell 34-41, 1991-2017

