Effekten af afdragsfrie realkreditlån på reale huspriser: En undersøgelse af Danmarks case

4. semester, cand.oecon

Gruppe 2

Andreas Methling, Kristoffer Herrig Thorndal

Aalborg Universitet

1. juni 2023

Titel:

Effekten af afdragsfrie realkreditlån på reale huspriser: En undersøgelse af Danmarks case

Tema:

Speciale

Projektgruppe:

2

Deltager(e):

Andreas Methling

Kristoffer Herrig Thorndal

Vejleder:

Mikael Randrup Byrialsen

Sidetal:

79

Antal normalsider:

56.25 sider

Afleveringsdato:

1. juni 2023

Abstract

This thesis investigates the implications of the implementation of interest-only mortgages in Denmark on real housing prices. Employing the synthetic Difference-in-Differences method the findings reveal a positive effect of the adoption of interest-only loans on the house price index in Denmark. The validity of these results is substantiated through an ARDL Bounds test, which confirms the statistically significant relationship between interest-only mortgages and the trajectory of Danish real estate prices. Specifically, a 1% increase in the proportion of interest-only mortgage loans is associated with a 0.23% long-term increase in Danish real estate prices. Furthermore, this study underscores the pivotal role played by policy interventions and decisions in shaping the dynamics of the housing market. It underscores the criticality of conducting meticulous assessments to accurately gauge the potential consequences of such policy measures in order to foster stability and sustainability within the housing market. Consequently, this study underscores the significance of interest-only mortgage loans as a policy instrument while acknowledging their dual impact on the housing market. Further research endeavors are recommended to comprehensively elucidate the multifaceted factors that underpin the relationship between interest-only mortgages in and real housing prices in Denmark.

Indholdsfortegnelse

1 Indledning **6**

1.1 Problemformulering 9

1.2 Udvidet problembestemmelse/afgrænsning 9

1.3 Projektmetode 10

2 Litteraturgennemgang **11**

3 Økonometrisk metode **15**

3.1 Dimensionality Reduction 15

3.1.1 Principal Component Analysis 15

3.1.2 Least Absolute Shrinkage and Selection Operator 17

3.2 Stationaritet 18

3.3 Augmented Dickey-Fuller 20

3.4 ARDL bounds test 22

3.5 Diagnostikker 24

3.6 Difference-in-Differences 27

3.7 Syntetisk Difference-in-Differences 29

3.7.1 Syntetisk kontrolmetode 29

3.7.2 Kombinationen af SCM og DiD 30

3.8 Data 30

4 Deskriptiv analyse **35**

4.1 Udvælgelse af variable 35

4.1.1 PCA 35

4.1.2 LASSO-regression 38

4.2 Udvælgelse af lande til Syntetisk DiD 39

5 Analyse **43**

5.1 Syntetisk Difference-in-Differences 43

5.1.1 SCM 43

5.1.2 SDiD 46
5.2 ARDL Bounds test 48
5.3 Robusthedstjek 52
5.4 OLS-estimering på filtreret data 55

6 Diskussion **58**
6.1 Validering af resultater 58
6.1.1 DiD-analysen 58
6.2 Fejlkorrektionsmodeller 60
6.3 Politisk perspektiv og relevans 62

7 Konklusion **65**

8 Litteratur **67**

Appendiks **73**
Appendiks A: Den overordnede fejlkorrektionsmodel 73
Appendiks B: Den korte fejlkorrektionsmodel 76
Appendiks C: OLS den cykliske del af variable 78
Appendiks D: Robusthedstjek af DiD estimater 79
Appendiks E: Andet 79

1 Indledning

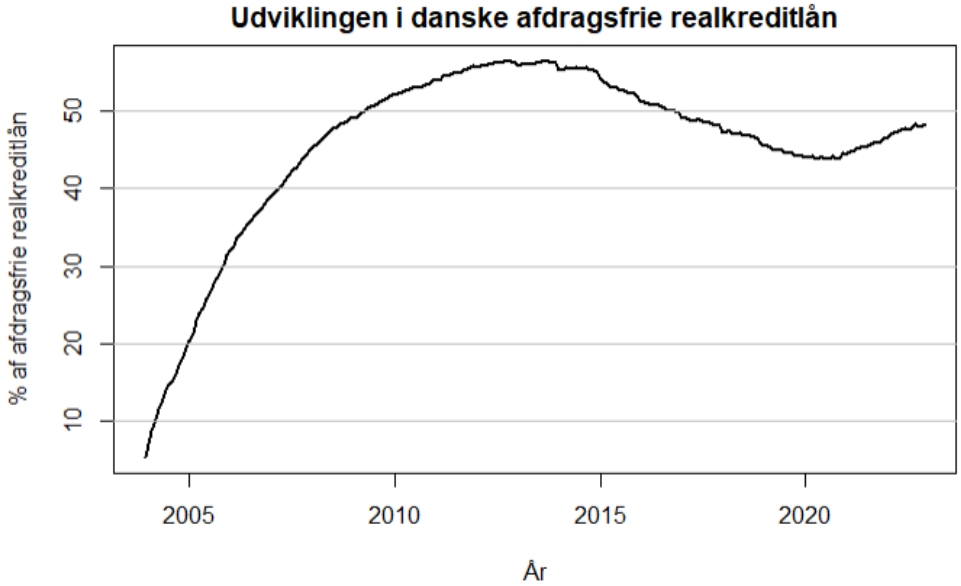
Den 1. oktober 2023 markerer en milepæl for det danske realkreditmarked, da den danske befolkning på denne dato har haft mulighed for at anskaffe sig afdragsfrie realkreditlån i 20 år. Dette 20 års jubilæum giver anledning til at evaluere indførslen af lånetypen på det danske realkreditmarked, og hvilken betydning den har haft for boligmarkedet i sin helhed.

Kort fortalt er et afdragsfrit realkreditlån et lån, hvor der i en periode, normalt kun op til 10 år af lånets løbetid, er mulighed for at sætte sine afdrag på pause og kun betale rente og bidrag. Fordelene herved vil være, at det nu er muligt at betale af på anden dyrere gæld i perioden uden afdrag eller, at der i de ti år uden afdrag vil være mere luft i økonomien og mulighed for at bruge sin indkomst her og nu til eventuelle renoveringsprojekter i forbindelse med huskøb eller andre umiddelbare udgifter (Totalkredit, 2023). Det er dog vigtigt at pointere, at lånet efter de 30 år stadig skal betales tilbage, og lånet derfor i sin helhed vil være omkostningstungt, da der betales større renter over lånets løbetid og en højere bidragssats til realkreditinstituttet end et tilsvarende lån med afdrag (Totalkredit, 2023).

Det blev ved lov tilladt for danske boligejere pr. 1. oktober 2003 at optage afdragsfrie realkreditlån. Afdragsfriheden blev indført med baggrund i et ønske fra politisk side om at give danskerne en mere fleksibel tilgang til deres boliglån både for unge såvel som ældre boligejere. En analyse fra 2007 lavet af Realkreditrådet (2007) viser, hvordan boligejere, der optog et afdragsfrit lån i 2007, anvendte denne fleksibilitet. Kun 28% af låntagerne brugte deres afdragsfrihed til at løfte deres forbrug, imens 18% har benyttet det til at håndtere lønnedgang, orlov, uddannelse, barsel, skilsmisse eller dødsfald. Den største gruppe brugte afdragsfriheden til at afdrage på anden dyrere gæld, øge deres opsparing/pension eller investere yderligere i boligen. Dette understreger fleksibiliteten af det afdragsfrie lån, og hvorfor det fra politisk side umiddelbart var fordelagtigt at indføre. (Økonomi - og erhvervsministeriet, 2008)

Siden sin indførsel i 2003 er de afdragsfrie realkreditlån blevet et af de foretrukne lån hos den danske befolkning. Udviklingen i andelen af realkreditlån, der har afdragsfrihed er vist i figur 1. Ud fra figuren kan det observeres, at et år efter loven trådte i kraft

var 17% af de udstedte realkreditlån afdragsfrie, og efter 5 år var låntypen allerede mere populær end alternativet med afdrag. Selvom populariteten af afdragsfrihed har set en mindre nedgang siden sin top i 2013 på 56.5%, så er lige over 48% af alle aktive danske realkreditlån stadig uden afdrag.

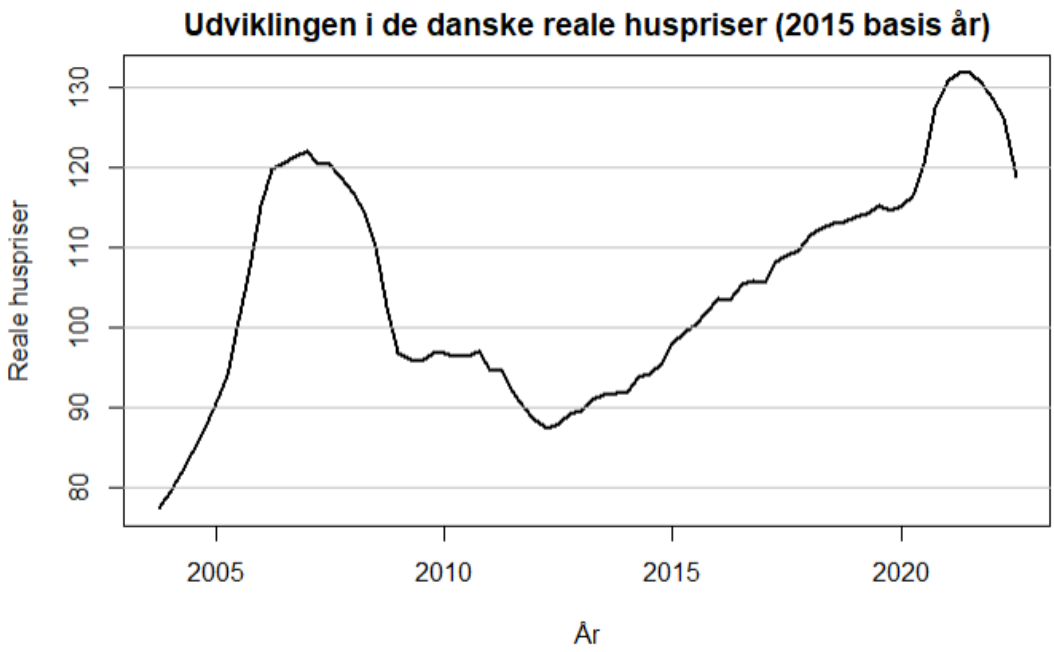


Figur 1: Udviklingen i andelen af afdragsfrie realkreditlån siden december 2003.
Kilde: (Nationalbankens Statistikbank, 2023) & (Nationalbankens Statistikbank, 2013a).

Ovenstående figur bevidner om, at danskerne har taget afdragsfriheden til sig og benytter den i et stort omfang. De positive aspekter af afdragsfriheden er til at få øje på for de danske husholdninger, men den førnævnte fleksibilitet som lånet tilbyder kan have overskygget de negative effekter, der kan være sværere for en gennemsnitlig husholdning at gennemskue.

Den pludselige og lettere tilgængelige adgang til kapital har haft den konsekvens, at boligpriserne har oplevet markante stigninger siden indførelsen af afdragsfriheden. Disse stigninger kan ses af figur 2. Selvom stigningerne i boligpriserne kan virke positive for de boligejere, der oplever dem, kan de i perioder med økonomisk krise og høj volatilitet have alvorlige konsekvenser. Dette skyldes, at mange husholdninger kan ende i en situation med teknisk insolvens, hvor de ikke er i stand til at tilbagebetale deres realkreditlån. På de fem år det tog afdragsfrie lån at blive danskernes foretrukne realkreditlån, steg huspriserne 24.9%, hvilket svarer til 5% hvert eneste år, og det er vel at mærke efter det store fald under finanskrisen. Ses der udelukkende på udviklingen fra fjerde kvartal 2003

til første kvartal 2007 steg de reale huspriser på tretten kvartaler med 36%, hvilket svarer til 2.77% per kvartal eller lidt over 10% om året.



Figur 2: Udviklingen i de danske reale huspriser siden tredje kvartal 2003.
Kilde: (OECD.data, 2022b).

Afdragsfriheden er ikke den eneste faktor, der har spillet ind på boligprisernes himmelflugt, hvor perioden efter finanskrisen især har været kendetegnet ved lave renter og høj efterspørgsel, hvilket også har haft en positiv effekt på boligpriserne. Det Systematisk Risikoråd har foreslået forskellige værktøjer til at lempe stigningerne og dermed dæmpe volatiliteten på boligmarkedet. Disse omfatter yderligere begrænsninger på udbredelsen af variabelt forrentede realkreditlån, øget krav til udbetalingsstørrelsen eller begrænsning af afdragsfriheden. (Rasmussen, 2021)

Ovenstående aspekter indikerer, at det kunne være værd at kigge nærmere på, om indførslen af muligheden for at optage realkreditlån med afdragsfrihed havde en reel effekt på boligpriserne, eller stigningerne er sket på grundlag af andre markedsfænomen i omverdenen. Givet de indledende overvejelser og observationer søger projektet svar på nedenstående problemformulering.

1.1 Problemformulering

Hvilken betydning har implementeringen af afdragsfrie realkreditlån i Danmark haft på de reale huspriser?

1.2 Udvidet problembestemmelse/afgrænsning

Der er en relativt moderat mængde akademisk litteratur, der analyserer årsagssammenhængen mellem indførslen af afdragsfrie realkreditlån og de reale huspriser både i et dansk og et internationalt perspektiv, hvorfor dette projekt forsøger at bidrage hertil. Den danske belysning rummer også nogle naturlige interesseforhold, hvorfor dette projekt er begrænset til den danske økonomi. Derudover betragter projektet en tidshorisont fra 1. januar 1994 til 31. december 2018.

Begrundelsen for at vælge netop dette tidsrum er primært baseret på tilgængeligheden af dansk makrodata. Desuden vil implementering af kriser i tidsserieanalyser i nogen grad forvrænge analysen i form af strukturelle brud, volatilitet og usikkerhed i dataet. Af denne grund inkluderer dataserierne ikke en nylig krise som COVID-19-krisen, og er derfor begrænset til at indeholde observationer frem til og med 2018.

Projektet fokuserer specifikt på priser for beboelsesejendomme som en samlet betegnelse for private husstande. I den økonometriske analyse anvendes der specifikt prismål for énfamiliehuse, da dette er datagrundlaget indsamlet af Danmarks Statistik.

Udvælgelsen af danske makrovariable, der indgår i projektets undersøgelse, er hovedsageligt baseret på en vurdering af relevans og anvendelighed. Fx vil implementeringen af inflation ikke være relevant, da der i projektet undersøges de reale priser, hvor inflation allerede er taget højde for. Yderligere vil et utal af makrovariable indgå, da variablene vil blive testet for deres anvendelighed, inden de inddrages i de økonometriske modeller.

Rentetilpasningslån vil i en senere litteraturgennemgang blive nævnt, men er ikke yderligere noget projektet specifikt vil anvende eller tage højde for, hvorfor der kun vil blive fokuseret på den samlede pulje af realkreditlån, med eller uden afdrag, uafhængigt af om de underliggende lån er med fast eller variabel rente.

Videre afgrænsning sker i form af metoder til projektet, hvor der i den eksisterende litteratur benyttes en række forskellige tests og modeller, som hver har sine styrker og

svagheder. I dette projekt sker valget af metoder i nogen grad ud fra tilgængelighed, hvor ARDL Bounds tests sammen med fejlkorrektionsmodeller udgør det primære vidensfundament givet fagindholdet for kurser på bachelor- samt kandidatuddannelsen i økonomi på Aalborg Universitet. Foruden bachelor- og kandidatuddannelsens fagindhold er metoder som Difference-in-Differences og Synthetic Control Method blevet anvendt. Disse er essentielle for undersøgelsen, da de giver et andet indblik i det undersøgte forhold end de mere traditionelle metoder.

1.3 Projektmetode

Projektets fremadrettede fremgangsmetode er opbygget således, at der i sektion 2 først redegøres for tidligere studiers resultater omhandlende den undersøgte problemstilling. I sektion 3 gennemgås de tests og modeller, der anvendes til at udføre analysen i projektet, samt de diagnostikker, der er med til at kvalitetssikre denne. Der gives her ligeledes en kort gennemgang og forklaring af de anvendte dataserier. Videre undersøges der i sektion 4 hvilke dataserier, der samlet bedst egner sig til at blive anvendt i en økonometrisk analyse vha. dimensionality reduction metoder, samt hvilke lande der egner sig bedst til DiD-analysen. Sektion 5 anvender de opstillede modeller og tests til at estimere og fastslå det langsigtede forhold mellem husholdningernes anvendelse af afdragsfrie realkreditlån og de reale huspriser. Hertil diskuteres analysens resultater i sektion 6 med udgangspunkt i tidskonsistente estimater, forskellige længder af dataserier og metodernes anvendelighed til at løse den opstillede problemstilling. Til sidst inkorporeres et politisk perspektiv til diskussionen af resultaterne.

2 Litteraturgennemgang

I sidste afsnit blev projektets problemformulering introduceret, hvorfor det nu er nærliggende at skabe en forforståelse på emnet. Således vil der i det følgende afsnit blive undersøgt, hvilke metoder og resultater den tidligere empiri har dannet på emnet. Dette skal ses som værende en vidensdannelse, der ligger til grund for at projektet kan udformes samt holde den røde tråd.

Flere artikler har undersøgt fluktueringer på boligpriser, mens færre fokuserer direkte på afdragsfrie realkreditlån og disses effekter på boligpriser. En relevant artikel der gør netop dette, er studiet af Dam et al. (2011), som i deres artikel "Boligboblen der bristede: Kan boligpriserne forklares? Og kan deres udsving dæmpes?" undersøger mulige forklaringsfaktorer, der kan spille ind på boligpriserne. Artiklen tager udgangspunkt i boligprisboomet i 00'erne og anvender hertil en økonometrisk analyse til at analysere hvilke faktorer, der har en effekt på boligpriser i Danmark. Undersøgelsen foretages af Danmarks Nationalbank på baggrund af dansk data fra 1999 til 2010. Der anvendes i artiklen flere dynamiske simulationer til at kortlægge opstillede variables præcise estimerede forklaringsgrad på udviklingen i boligpriser. Herunder opstilles en ny kontantprisrelation, da tidligere makroøkonomiske modeller ikke har kunnet forklare udsving i boligpriserne i årene op mod finanskrisen. I den nye kontantprisrelation indgår et vejet gennemsnit af hhv. brugerprisen (forklaret ved den forventede omkostning ved at eje en bolig) (Dam et al., 2011, s. 50) og den laveste førsteårsydelse muliggjort af de låneformer, der er tilgængelige på nuværende tidspunkt. Førsteårsydelsen fanger effekten af afdragsfrie realkreditlån på boligpriserne.

Kontantprisrelationen fanger ikke hele effekten af boligboblen, men det vurderes i artiklen, at den opfanger den samlede stigning i årtiet. Derfor anvendes den videre til at forklare de forskellige låneformers effekt på udviklingen i boligpriser. Her ses det fra kontantprisrelationen, at udviklingen i boligpriserne fra første kvartal 1999 til første kvartal 2007 uden implementering af de nye låneformer i 2003 havde været 40% fremfor 71%, som det reelt var tilfældet. Det tilføjes, at rentetilpasningslån og de afdragsfrie realkreditlån her tillægges lige stor betydning for boligpriserne. Et af resultaterne opstillet i artiklen argumenterer derfor også for, at fluktueringen i priser på boligmarkedet kan reduceres

ved at mindske adgangen til de afdragsfrie realkreditlån (Dam et al., 2011, s. 47).

Bäckman og Lutz (2020) har undersøgt indvirkningen af afdragsfrie realkreditlån på tilgængeligheden af boligkøb for de nederste sociale lag i Danmark. De fremlægger statistikker, der viser, at implementeringen af afdragsfrie realkreditlån ikke havde den ønskede effekt på tilgængeligheden, og at lånet i stedet blev anvendt af alle typer af boligejere og øgede omsætningshastigheden og især priserne på boliger i Danmark. Dette skyldes i deres optik, at udbuddet af boliger på kort sigt er fast og en lempelse af finansieringsmuligheder derfor giver øgede priser.

I deres analyse tester Bäckman og Lutz (2020) betydningen af fire variable på procentdelen af realkreditlån, der er afdragsfrie i hver af de danske 98 kommuner aggregeret ud fra mikrodata fra Danmarks Statistik. De anvender de gennemsnitslige boligpriser for hver enkelt kommune, den procentvise kvadratmeter prisstigning fra 1998-2002 og fra 2002-2007 i hver kommune og til sidst den gennemsnitlige indkomststigning i hver enkelt kommune. Deres resultater viser, at alle variable er signifikante, hvis de anvendes i en samlet regression, og at koefficienten på variablen for boligpriser er tre gange større end de andre. De konkluderer derfor, at boligpriser er den største drivkraft til at forudsige anvendelsen af afdragsfrie realkreditlån på kommunalt niveau i Danmark.

Yderligere konkluderes, at mængden af boligejere i alle sociale lag forblev den samme over analyseperioden, at karakteristika for nye boligejere forblev de samme og dermed ikke hjalp de mere udsatte danskere og at boligpriserne steg på grund af højere efterspørgsel, hvilket på trods af lempelser af finansieringsmuligheder gjorde nettoomkostningerne ved et boligkøb stort set uændret.

Karpestam og Johansson (2019) har undersøgt forholdet imellem introduktionen af afdragsfrie realkreditlån i Danmark i 2003 og de reale boligpriser, antal solgte boliger og forholdet mellem husholdningernes realkreditgæld ift. disponibel indkomst. Deres analyse viser, at introduktionen af afdragsfrie realkreditlån havde en effekt både på den store stigning i reale huspriser og stigningen i realkreditgæld op igennem 00'erne. Dog viste deres resultater ikke, at låntypen har haft en signifikant indvirkning på mængden af solgte boliger.

Karpestam og Johansson (2019) sætter spørgsmålstegn ved, om afdragsfrie realkre-

ditlån er det rigtige redskab til at sikre tilgængeligheden af boligkøb for alle husholdningstyper, da størstedelen af de danske husholdninger muligvis ikke har forståelse for den øgede risiko de påtager sig, ved at optage en større gæld som resultat af afdragsfriheden på deres lån. Yderligere fremlægger Karpestam og Johansson (2019) deres mest konservative estimater, som viser, at afdragsfrie realkreditlån har øget de reale huspriser og husholdningernes realkreditgæld med hhv. 20 og 10 % imellem 2004 og 2013. Derfor vil det være problematisk for politiske beslutningstagere at få gældsniveauet tilbage til niveauet fra før 2003, da det kan have store konsekvenser for den samlede efterspørgsel.

Afdragsfrie realkreditlån er ikke kun noget, der har interesseret danske forskere - Barlevy og Fisher (2021) har, i stedet for at undersøge Danmark, undersøgt brugen af afdragsfrie realkreditlån i USA og effekten heraf under bolig boomet i 00'erne. Deres undersøgelse tager udgangspunkt i, at afdragsfrie realkreditlån var mere anvendt i de områder, der så de største prisstigninger, hvilket de forsøger at finde årsagerne til. Deres analyse viser, at afdragsfrie realkreditlån havde en signifikant betydning for stigningerne i boligpriserne i amerikanske byer med et uelastisk udbud af boliger. Desuden steg brugen af afdragsfrie realkreditlån og boligpriserne mere i områder, der også så store indkomststigninger. Dette passer ifølge Barlevy og Fisher (2021) godt ind i opbygningen af afdragsfrie realkreditlån, hvor låntager har mulighed for at have en lavere indkomst indledningsvist, da der kun skal betales renter på lånet. Med tiden, når afdragsfriheden ophører, er låntagerens indkomst steget og denne har dermed bedre mulighed for at betale renter og afdrag på lånet.

Yderligere finder de, at realkreditlån uden afdrag opildner eller i hvert fald får flere til at spekulere i boliger, fordi en investor kun vil skulle betale renter og ikke afdrag i den tid, personen regner med at holde boligen i sin portefølje. Problematikkerne herved omfatter, at flere spekulanter har et ønske om at handle boliger, og dermed driver det priserne op, og det gør det svært for især unge at komme ind på boligmarkedet specielt i de større byer. Deres løsning herpå ville være at gøre det muligt for kreditorer at kunne forfølge debitor for yderligere aktiver, hvis gælden overstiger husets værdi.

Fra litteraturen er det tydeligt for såvel de danske som den amerikanske artikel, at de afdragsfrie realkreditlån har haft en ikke-negligerbar effekt på udviklingen i boligpriser. Magnituden heraf er dog stadig usikker, mens empirien er mangelfuld. Derfor vil det

fremtidige forløb i projektet være uændret.

Der vil i næste afsnit blive præsenteret essentielle modeller og metoder brugt i dette projekt - både metoder anvendt i projektet og metoder, som vurderes vigtige for forståelsen af de anvendte metoder.

3 Økonometrisk metode

Den økonometriske metode anvendt i projektet vil nu blive uddybet. Først vil der gives en kort præsentation af dimensionality reduction. Derefter vil stationaritet ved tidsserier blive forklaret. Hernæst vil der gives en kort forklaring af de anvendte økonometriske modeller i projektet, samt metoden benyttet til at sikre overholdelse af OLS-antagelserne i de anvendte modeller. Til sidst vil de anvendte dataserier kort præsenteres samt uddybning af eventuelle anvendte transformationer af disse.

3.1 Dimensionality Reduction

Dimensionality reduction analyser anvendes bl.a. for at undgå overfitting af dataet, og da det er et vigtigt redskab til at hjælpe mod multikollinearitet. En af de vigtigste grunde hertil er for at formindske "The Curse of Dimensionality". Dette koncept bygger på den idé, at jo flere features (variable) et datasæt indeholder, jo større antal observationer er nødvendigt for at finde sammenhænge i dataet. Begrebet Curse of Dimensionality er oftest nævnt i sammenhæng med Machine Learning, hvor flere dimensioner betyder progressivt mere træningstid for de opbyggede modeller, men begrebet kan også anvendes indenfor økonometri. Hvis mængden af forklarende variable stiger, bliver det således eksponentielt sværere at finde skjulte sammenhænge bag dataet (Yiu, 2019).

Udfordringerne ved dimensionalitet kan mindskes ved anvendelse af dimensionality reduction metoder. Herunder findes både feature extraction og feature selection. Dette projekt begrænser sig til feature extraction. Et eksempel herpå er Principal Component Analysis, som uddybes i næste afsnit.

3.1.1 Principal Component Analysis

En Principal Component Analysis, eller en PCA, er en dimensionality reduction metode, der anvendes indenfor unsupervised machine learning som et redskab til at udvælge variable og vælge at beholde kun den vigtigste information fra et datasæt (Abdi og Williams, 2010, s. 433). Dette muliggør, at vi kan skabe en model med færre features, men stadig beholde mest muligt af forklaringsgraden. Metoden inddeler inputvariablene i principal components. Antallet af PC's vælges først, hvorefter testen giver hver variabel en talværdi

tilhørende hver PC. Variablene med de højeste absolutte værdier er de vigtigste variable til at forklare den tilhørende PC.

Fremgangsmetoden i en PCA kan teoretisk inddeles i 5 trin (Keita, 2023):

1. Normalisering

Da hver variabel kan være meget forskellig i magnitude og derfor have meget forskellige varianser, startes der med at skalere datasættet, så alle variable er på samme niveau. I R kan dette opnås ved at anvende scale()-funktionen, der først udfører centrering og dernæst skalering. Centrering justerer datasættet ved at trække kolonnemiddelværdien fra hver værdi i kolonnen, og skalering sikrer, at alle kolonner har samme varians. Dermed får hver variabel en middelværdi på 0 og en standardafvigelse på 1. Dette resulterer i data, der potentielt reducerer bias mellem variablene.

2. Kovarians-matrice (eller Korrelation-matrice)

Efter normaliseringen opstilles og defineres en kovarians-matrix eller en korrelationsmatrix.

3. Egenvektorer og egenværdier

I dette trin udregnes egenvektorer og egenværdier ud fra den opstillede kovarians- eller korrelationsmatrix. Egenvektoren bestemmer hvilken retning i den pågældende dimension, som variablen har, mens egenværdien bestemmer længden på denne vektor. Egenværdien bestemmes ud fra størrelsen på variansen af dataet. Således kan dette par udgøre et mål for, hvor vigtig variablen er indenfor den givne dimension. Egenvektoren med den største egenvektor bliver til den første (og vigtigste) PC, mens egenvektoren med den andenhøjeste egenværdi bliver til den anden PC.

4. Udvælgelse af Principal Components

Der findes ikke én metode til at bestemme, hvor meget varians der skal forklares (og dermed hvor mange PC's der vælges), men nogle mulige metoder er:

1. Sæt en tærskel på hvor meget kumuleret varians man er tilfreds med at have forklaret ved sine PC's.

2. Lav Cattell's scree-test (også kaldt albuetesten), hvor man vha. et scree-plot tester,

hvornår grafen for varians i de forskellige dimensioner går fra stejl til flad, og derfra vælger det antal komponenter, som ligger til venstre for "albuen". Se figur 4 for illustration af et scree plot. (Oxford Reference)

3. Brug Kaiser's regel ved at finde egenværdierne tilhørende komponenterne og vælg alle PC's over 1 i egenværdi.

I dette projekt er det primært Kaiser's regel der lægges vægt på i udvælgelsen af PC's, da denne specificerer et præcist beregnet mål, og ikke er bygget på subjektiv vurdering. De andre metoder vil dog også blive fremvist, som supplerende begrundelse for valget.

5. Data transformation i et nyt dimensionelt rum

Til sidst transformeres dataet til at blive vist i det nye antal dimensioner. Dette visualiseres fx gennem et todimensionelt plot, hvor x-aksen er første dimension og y-aksen er anden dimension.

Der er i en PCA nogle nødvendige antagelser for at kunne opnå nogle brugbare resultater (Keita, 2023). Den første er, at der findes en lineær sammenhæng mellem de undersøgte variable. Derudover antager PCA, at de PC's, som indeholder størst varians er de vigtigste komponenter. Grunden til at den store varians pålægges stor vigtighed skal ses som, at variable med stor varians indeholder stor forklaringskraft; en helt flad lineær tidsserie uden udsving indeholder ikke meget information, hvorimod en svingende tidsserie kan forsøges forklaret eller bruges bedre til at forklare andre tidsserier. Det samme gør sig gældende for PC's, som er lineære kombinationer af inputvariablene.
Dette nødvendiggør, at der tages højde for, hvorvidt datasættet indeholder store outliers, da et stort antal outliers kan skabe bias i resultaterne. Derudover præsterer PCA bedst ved store datasæt.
Det skal nævnes, at ulig flere andre dimensionality reduction metoder, så er PCA deterministisk, hvilket betyder, at samme udførte analyse skaber samme resultater hver gang analysen udføres.

3.1.2 Least Absolute Shrinkage and Selection Operator

Least Absolute Shrinkage and Selection Operator (Lasso) regressionen er en type lineær regression anvendt indenfor machine learning. Metoden anvender en såkaldt shrinkage-

teknik kaldet L1-regulering (Glen, a). Denne type regulering bruges, når formålet er at eliminere variable, hvorimod L2-regulering anvendes til at mindske koefficienterne (uden at fjerne nogle variable) (Glen, b). L1-reguleringen fremmer simple modeller ved at tilføje en straf, der svarer til den absolutte værdi af koefficienternes størrelse. Denne straf kan føre til, at nogle af variablenes koefficienter får en nulværdi. Det er disse variable med nulværdier, som kan fravælges i variabeludvælgelsen for derigennem at skabe en mere simpel model. For at bestemme styrken på straffen anvender LASSO regressionen en tuning-parameter kaldet λ. En større værdi af λ fører til mere shrinkage, større bias, men lavere varians, mens en mindre værdi af λ reducerer shrinkage, hvilket resulterer i lavere bias, men potentielt højere varians. Dette er et trade-off, som λ parameteren udregner. Dermed kan det siges, at i takt med at λ stiger styrkes straffen og dermed opstår der flere nulværdier i variablenes koefficienter.

Processen anvender krydsvalidering, hvilket indebærer at opdele datasættet i to dele: et træningssæt og et valideringssæt. Træningssættet bruges til at træne modellen, mens valideringssættet bruges til at evaluere på modellens ydeevne. LASSO-regressionen anvender k-fold krydsvalidering, som opdeler datasættet i k mindre dele, eller såkaldte folds. Modellen trænes k gange, hvor hver gang bruger $k - 1$ folds til træning og den tilbageværende fold til validering. Dette gentages, indtil hver fold har fungeret som valideringsfold (Brownlee, 2020). Dette leder ud i en λ-værdi, som således styrer, hvilke variable der i sidste ende bliver elimineret.

3.2 Stationaritet

I arbejdet med økonomiske tidsserier er det essentielt at fastslå, hvornår de anvendte serier er stationære, da unit roots i tidsserierne kan gøre dem tidsafhængige og dermed skabe uforudsigelighed, hvilket gør serierne uanvendelige i økonometrisk modelarbejde. En tidsserie siges at være stationær, når der ingen systematisk ændring er i middelværdien og dermed ingen trend, og når der ingen systematiske forstyrrelser er i variansen. Med andre ord skal seriens middelværdi og varians være tidsinvariant og dermed ikke afhænge af tid. Ovenstående betingelser kaldes svag stationaritet og er den form for stationaritet, der vil

blive anvendt i dette projekt. Svag stationaritet bygger på nedenstående tre antagelser:

$$E[y_t] = \mu < \infty \tag{3.1}$$

$$Var[y_t] = E[(y_t - \mu)^2] = \gamma_0 < \infty \tag{3.2}$$

$$Cov(y_t, y_{t-k}) = E[(y_t - \mu) * (y_{t-k} - \mu)] = \gamma_k,\ For\ k = 1, 2, 3, ... \tag{3.3}$$

Ligning 3.1 viser, at tidsserien har et konstant gennemsnit, μ, der dermed ikke afhænger af tiden. De resterende to ligninger 3.2 og 3.3 viser, at varians-kovarians matricen ligeledes er konstant og tidsinvariant.(Pfaff, 2008)

Tidsserier hvis middelværdi og varians afhænger af tid kan gøres stationære ved at tage første difference af ligningen. Dette vil blive vist med et eksempel på en random walk med drift:

$$\begin{aligned} Y_t &= \mu + Y_{t-1} + \varepsilon_t \\ hvor\ \varepsilon_t &\sim IID(0, \sigma^2) \end{aligned} \tag{3.4}$$

Denne series middelværdi og varians er ikke konstant og afhænger dermed af tid, som det fremgår af ligning 3.5 og 3.6:

$$\begin{aligned} E\{Y_t\} &= \mu + Y_{t-1} + \varepsilon_t \\ &= \mu + \mu + Y_{t-2} + \varepsilon_{t-1} + \varepsilon \\ &= Y_0 + \sum_{i=1}^{t} \mu + \sum_{i=1}^{t} \varepsilon_t \\ &= E\{Y_0 + \sum_{i=1}^{t} \mu + \sum_{i=1}^{t} \varepsilon_t\} = Y_0 + t\mu \end{aligned} \tag{3.5}$$

$$\begin{aligned} Var\{Y_t\} &= Var\{Y_0 + \sum_{i=1}^{t} \mu + \sum_{i=1}^{t} \varepsilon_t\} \\ &= 0 + 0 + \sum_{i=1}^{t} Var\{\varepsilon_t\} \\ &= t\sigma^2 \end{aligned} \tag{3.6}$$

Tages første difference af serien fjernes tidskonsistensen og serien bliver dermed stationær:

$$\begin{aligned} Y_t - Y_{t-1} &= \mu + Y_{t-1} - Y_{t-1} + \varepsilon_t \\ \Delta Y_t &= \mu + \varepsilon_t \\ hvor\, \varepsilon_t &\sim IID(0, \sigma^2) \end{aligned} \tag{3.7}$$

Den opnåede serie i ligning 3.7 har en konstant middelværdi, μ, og varians, σ^2, og kan kaldes en stationær proces af første difference. Ovenstående har vist, at tidsserier kan gøres stationære ved at tage d antal differencer af serien med d antal unit roots. (Enders, 2014)

For at finde integrationsordenen af de i projektet anvendte tidsserier vil augmented Dickey-Fuller testen anvendes, der vil blive beskrevet nærmere i næste undersektion.

3.3 Augmented Dickey-Fuller

Bestemmelsen af integrationsordenen for de i projektet anvendte tidsserier er et vigtigt element. Først og fremmest fordi unit roots som tidligere nævnt kan skabe uforudsigelighed i serien, men også fordi fremtidige metoder/modeller i visse tilfælde stiller krav til seriernes integrationsorden.

Dickey og Fuller (1979) betragtede tre forskellige regressioner i deres undersøgelse af test for unit root. Den første er en "pure random walk", den næste tilføjer et intercept også kaldet et drift term og den sidste tilføjer både et drift term og en lineær trend. De opstillede ligninger fremgår af ligning 3.8 til 3.10.

$$Y_t = \beta_1 Y_{t-1} + \varepsilon_t \tag{3.8}$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \tag{3.9}$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 t + \varepsilon_t \tag{3.10}$$

Dickey-Fuller ligningerne opstår, når Y_{t-1} fratrækkes på begge sider af lighedstegnet:

$$\Delta Y_t = \gamma Y_{t-1} + \varepsilon_t \tag{3.11}$$

$$\Delta Y_t = \beta_0 + \gamma Y_{t-1} + \varepsilon_t \tag{3.12}$$

$$\Delta Y_t = \beta_0 + \gamma Y_{t-1} + \beta_2 t + \varepsilon_t \tag{3.13}$$

Den interessante parameter i disse ligninger er γ, der også kan skrives som $\gamma = (\beta_1 - 1)$. I tilfælde af, at $\gamma = 0$, hvilket også vil sige, at $\beta_1 = 1$, så vil serien indeholde en unit root. Testen foretages ved at estimere en af ovenstående ligninger med Ordinary Least Squares (OLS) for at opnå den estimerede værdi af γ og den tilhørende standardfejl. Dernæst sammenligner man den resulterende t-statistik med de korrekte kritiske værdier fremlagt af Dickey og Fuller (1979), da t-værdien under nulhypotesen ikke følger en normal t-distribution. Hvis nulhypotesen ikke afvises lider modellen af unit root, hvorved antagelsen om svag stationaritet ikke vil være overholdt. Derfor tages den første difference af variablen, og testen foretages på ny. Hvis nulhypotesen hernæst afvises vil variablen være svagt stationær og variablen være integreret af første orden $I(1)$.

Fremgangsmåden er den samme uafhængigt af hvilken af de tre ligninger, der bliver anvendt, dog vil de kritiske værdier for t-statistikken ændre sig ved at inkludere et intercept og/eller en trend i testen.

Ligning 3.11-3.13 kan også omskrives til autoregressive processer, hvor de samme kritiske værdier anvendes til at teste for unit root:

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta Y_{t-i+1} + \varepsilon_t \tag{3.14}$$

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta Y_{t-i+1} + \varepsilon_t \tag{3.15}$$

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \alpha_2 t + \sum_{i=2}^{p} \beta_i \Delta Y_{t-i+1} + \varepsilon_t \tag{3.16}$$

Dickey-Fuller (DF) tests hvor der inkluderes lags af den afhængige variabel kaldes også augmented Dickey-Fuller (ADF) test og tester ligeledes hypotesen om at $gamma = 0$. Omskrivning fra DF-test til ADF-test sker på baggrund af, at DF-testen antager at ε_t og ε_{t-k} er ukorreleret for alle $k > 0$, hvis dette ikke er tilfældet, er der seriekorrelation i fejlleddet. Seriekorrelation i fejlleddet indebærer, at der er en systematisk bevægelse i fejlleddet, som modellen ikke opfanger ved de inkluderede variable, hvilket vil påvirke standard fejl og variansen i modellen. I sådanne situationer vil ADF-test foretrækkes, da

tilføjelse af lags af den afhængige variabel til højre side af ligningen i mange tilfælde vil fjerne seriekorrelationen. (Enders, 2014)

3.4 ARDL bounds test

Pesaran et al. (2001) udviklede i 2001 en ny tilgang til at teste for lineære sammenhænge mellem en afhængig variabel og et sæt af uafhængige variable. I modsætning til Engle-Granger tilgangen til test for kointegration, der beror sig på at regressere variable integreret af samme orden, er dette ikke nødvendigt i ARDL Bounds test. Dog skal de være $I(0)$ eller $I(1)$.

Betragt følgende model med to uafhængige variable, en konstant og et fejlled:

$$Y_t = \mu + \phi_1 X_{1t} + \phi_2 X_{2t} + \varepsilon_t \tag{3.17}$$

Modellen omskrives til en fejlkorrektionsmodel (ECM) som følger:

$$\begin{aligned}
Y_t &= \mu + \phi_1 X_{1t} + \phi_2 X_{2t} + \varepsilon_t \\
Y_t - Y_{t-1} &= -Y_{t-1} + \mu + \phi_1 X_{1t} + \phi_2 X_{2t} + \varepsilon_t \\
\Delta Y_t &= -Y_{t-1} + \mu + \phi_1 X_{1t} + \phi_2 X_{2t} + \varepsilon_t \\
\Delta Y_t &= -Y_{t-1} + \mu + [\phi_1 X_{1t}(-\phi_1 X_{1t-1} + \phi_1 X_{1t-1})] + \phi_2 X_{2t} + \varepsilon_t \\
\Delta Y_t &= -Y_{t-1} + \mu + [\phi_1 X_{1t} - \phi_1 X_{1t-1}] + \phi_1 X_{1t-1} + \phi_2 X_{2t} + \varepsilon_t \\
\Delta Y_t &= -Y_{t-1} + \mu + \phi_1 \Delta X_{1t} + \phi_1 X_{1t-1} + \phi_2 X_{2t} + \varepsilon_t \\
\Delta Y_t &= -Y_{t-1} + \mu + \phi_1 \Delta X_{1t} + \phi_1 X_{1t-1} + [\phi_2 X_{2t}(-\phi_2 X_{2t-1} + \phi_2 X_{2t-1})] + \varepsilon_t \\
\Delta Y_t &= \phi_1 \Delta X_{1t} + \phi_2 \Delta X_{2t} - Y_{t-1} + \mu + \phi_1 X_{1t-1} + \phi_2 X_{2t-1} + \varepsilon_t \\
\Delta Y_t &= \phi_1 \Delta X_{1t} + \phi_2 \Delta X_{2t} - [Y_{t-1} - \mu - \phi_1 X_{1t-1} - \phi_2 X_{2t-1}] + \varepsilon_t
\end{aligned} \tag{3.18}$$

ECM'en består af første differencer af de uafhængige variable og fejlkorrektionsleddet i parentesen. Udtrykket i parentesen svarer til modellens fejlled i levels i forrige periode, deraf navnet fejlkorrektionsmodel.

For at vise hvordan ARDL Bounds testen virker, sættes $\theta_1 = -1$ og er dermed koefficienten for Y_{t-1}, $\gamma_1 = -\frac{\phi_1}{\theta_1}$ og $\gamma_2 = -\frac{\phi_2}{\theta_1}$. Simplificeringerne kan ses af ligning 3.19.

$$\begin{aligned}
\Delta Y_t &= \phi_1 \Delta X_{1t} + \phi_2 \Delta X_{2t} + \theta_1 [Y_{t-1} - \mu - \phi_1 X_{1t-1} - \phi_2 X_{2t-1}] + \varepsilon_t \\
\Delta Y_t &= \phi_1 \Delta X_{1t} + \phi_2 \Delta X_{2t} + \theta_1 Y_{t-1} - \frac{\mu}{\theta_1} - \frac{\phi_1}{\theta_1} X_{1t-1} - \frac{\phi_2}{\theta_1} X_{2t-1} + \varepsilon_t \\
\Delta Y_t &= \phi_1 \Delta X_{1t} + \phi_2 \Delta X_{2t} + \theta_1 Y_{t-1} - \frac{\mu}{\theta_1} - \gamma_1 X_{1t-1} - \gamma_2 X_{2t-1} + \varepsilon_t
\end{aligned} \tag{3.19}$$

Da alle anvendte variable højest er integreret af første orden, $I(1)$, er den afhængige variabel på venstre side af ligningen med sikkerhed stationær. For at højresiden af ligningen ligeledes er stationær skal udtrykkene tilsammen være stationære. Derfor, for at der kan eksistere en lineær sammenhæng, skal en kombination af niveauvariablene være stationære. Hvis der ikke findes en kombination af niveauvariablene, der er stationære, er den eneste måde hvorpå højresiden af ligning kan være stationær ved, at koefficienterne på niveauvariablene alle er nul, og dermed eksisterer der ingen lineær sammenhæng. (Eriksen, 2020)

Nulhypotesen i en ARDL Bounds test er, at der ikke findes en langsigtet sammenhæng imellem variablene. Dette medgiver, at koefficienterne af de laggede niveauvariable i ECM'en tilsammen er nul:

$$\begin{aligned}
&H_0 : \theta_1 = \gamma_1 = \gamma_2 = 0 \\
&H_1 : \theta_1 \neq 0, \; eller \; \gamma_1 \neq 0, \; eller \; \gamma_2 \neq 0
\end{aligned} \tag{3.20}$$

Pesaran et al. (2001) giver kritiske værdier for F-statistikken, der kan bruges til at teste nulhypotesen. Disse værdier er udledt efter diverse antagelser. Den første er hvorvidt den sande ECM indeholder et intercept, en trend eller begge dele, og hvor mange variable, k, af laggede niveauvariable andre end Y_{t-1}, der indgår i niveau ligningen. Den anden antagelse omhandler fordelingen af F-statistikken på to måder. For det første, ved at øge antallet af uafhængige variable øges testens styrke, da antallet af frihedsgrader stiger. For det andet, fordi en ikke-standard distribution forekommer, når variable er $I(1)$ og antallet af mulige langsigtede sammenhænge, der kan eksistere mellem Y_t og alle x_i variablene stiger, som der bliver tilføjet flere $I(1)$ x_i variable, der kan kointegrere i blandt dem selv, hvilket tilstedekommer muligheden for flere potentielle langsigtet forhold.

Bounds testens to grænser er skabt ved at anvende to ekstreme tilfælde. Først hvor alle x_i variablene er $I(0)$. Dernæst hvor alle x_i variablene er $I(1)$. Disse to tilfælde danner

udgangspunkt for ARDL Bounds testens kritiske værdier, som test-statistikken skal sammenlignes med.

1. Hvis $\hat{F}$, test-statistikken skabt under H_0, er under den nedre grænse kan H_0 ikke afvises, og dermed kan det ikke afvises, at der ikke findes en langsigtet sammenhæng imellem Y_t og de anvendte uafhængige variable.
2. Hvis $\hat{F}$ er over den øvre grænse kan H_0 afvises, og dermed kan det afvises, at der ikke findes en langsigtet sammenhæng imellem Y_t og de anvendte uafhængige variable.
3. Hvis $\hat{F}$ falder imellem de to grænser er testen ubestemmelig.

Ovenstående framework vil blive anvendt i dette projekt til at finde frem til, om afdragsfrie realkreditlån har en langsigtet effekt på danske huspriser.

3.5 Diagnostikker

Det i opgaven anvendte datasæt, som er beskrevet i afsnit 3.8 indeholder 100 observationer. Grundet denne forholdsvis lave mængde observationer antages det ikke, at de anvendte modeller udviser asymptotiske egenskaber, og der er derfor blevet foretaget tests for at sikre, at de anvendte estimater er unbiased og de klassiske OLS-antagelser overholdes. Følgende undersektion vil nu beskrive disse antagelser og forklare, hvordan de er blevet taget højde for[1]. Følgende afsnit vil tage udgangspunkt i Wooldridge (2018) kapitel 10.

TS.1 Lineære parametre

Den første antagelse fastslår, at modellen skal være lineær i dens parametre, og dermed må der ikke indgå polynomier af nogen grad i modellens parametre. Dette er blevet testet ved hjælp af RESET testen, der undersøger, om modellens fittede værdier har et signifikant ikke lineært forhold. Denne antagelse er overholdt i den anvendte model. Resultatet af RESET testen kan ses i appendiks figur A.1.

TS.2 Ingen perfekt korrelation

Denne antagelse fastslår, at den enkelte forklarende variabel ikke må være en eksakt lineær

[1] Det er muligt at fremsende alle de udførte diagnostikker efter forespørgsel.

kombination af en af de andre forklarende variable. Korrelation mellem dem er tilladt, der må bare ikke være tale om perfekt korrelation, da modellen dermed ikke vil kunne estimeres ved hjælp af OLS. Dette er blevet testet for ved hjælp af en krydskorrelations test, der tester korrelationen imellem de anvendte variable. Der er blevet sat en øvre grænse på 0.85, og hvis variablene ligger over denne grænse kan de ikke inkluderes i den samme model. Resultater af krydskorrelations test kan ses af tabel A.2 i appendiks.

TS.3 Eksogenitet

Denne antagelse indebærer, at fejlleddet i tidspunktet t, u_t, er lig med 0 og ikke er korreleret med nogle af de forklarende variable i alle tidsperioder. Det vil sige, at fejlleddet i den anvendte model ikke må afhænge af nogle af de uafhængige variable. Dette kan omstilles matematisk således:

$$E(u_t|\mathbf{X}) = 0,\ t = (1, 2, 3, ..., n) \tag{3.21}$$

Når denne antagelse er overholdt er estimaterne strengt eksogene. Hvis antagelsen ikke er overholdt er der endogenitet i modellen, hvilket kan skyldes målefejl, der kan opstå ved omregning eller ændring af de oprindelige variable anvendt i projektet. En anden årsag kunne være "omitted variable bias", hvor en relevant variabel for modellen er udeladt og derfor kan skabe korrelation mellem fejlleddet og en af de forklarende variable. De fittede værdier for den anvendte model er plottet imod fejlleddet for at teste den antagelse. Grafen kan ses i appendiks figur A.1 og viser ingen tegn på problemer med systematiske udsving i fejlleddet.

TS.4 Homoskedasticitet

Denne antagelse indebærer, at variansen af fejlleddet u_t er konstant over tid, og uafhængig af de forklarende variable. Ved brud på denne antagelse siges fejlleddet at være heteroskedastisk. I dette projekt benyttes ARCH test til at undersøge, hvorvidt variansen er konstant over tid. Derefter anvendes Breusch-Pagan (BP) testen til at teste, om variansen er konstant ved ændringer i de forklarende variable. Begge tests viser, at den anvendte model er homoskedastisk, hvilket betyder, at modellens varians er konstant over tid og er uafhængig af de forklarende variable anvendt i modellen. Resultaterne af testene kan ses i appendiks figur A.2 og tabel A.3.

TS.5 Seriekorrelation

Denne antagelse omhandler seriekorrelation, som betyder, at der ingen korrelation må være mellem selve fejlleddet og lags af fejlleddet.

$$corr(u_t, u_s) = 0, \; for\, all\, t \neq s \tag{3.22}$$

Dette testes ved Ljung-Box testen i R. Denne antagelse er ligeledes overholdt i den anvendte model, og resultatet af testen kan ses i appendiks tabel A.4.

Efter antagelserne TS.1 til TS.3 er opfyldt er OLS estimaterne unbiased givet de forklarende variable. Efter indførelsen af TS.4 og TS.5, kan OLS estimaterne betegnes som BLUE (Best Linear Unbiased Estimates) givet de forklarende variable.

TS.6 Normalitet

Denne antagelse indebærer, at fejlleddet u_t er uafhængig af de forklarende variable og er IID (independently and identically distributed) med en middelværdi på 0 og en standardafvigelse på σ^2, hvilket kan skrives som: $Normal(0, \sigma^2)$. Dette medfører, at fejlleddet i modellen udvikler sig tilfældigt.

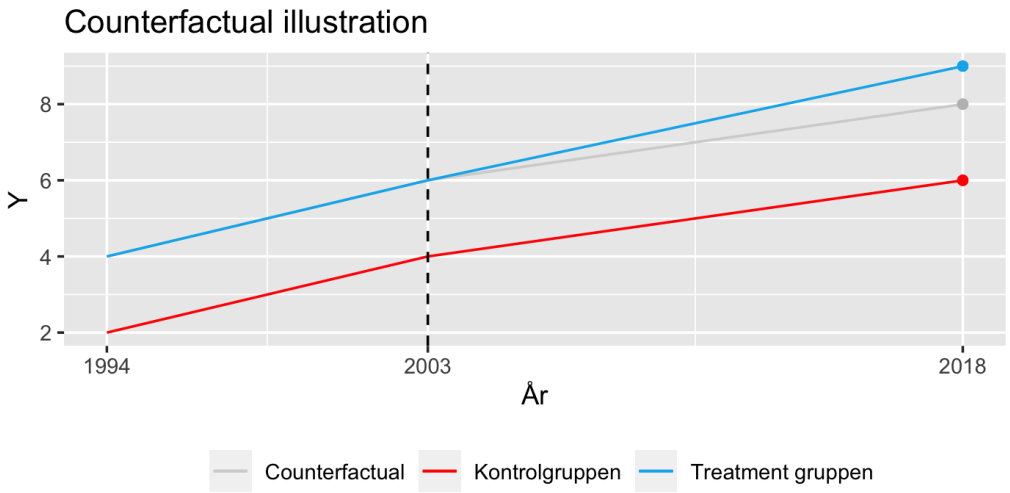
Ved overholdelse af TS.1 til TS.6 kan der nu anvendes t-statistik til at teste for statistisk signifikans ved individuelle forklarende variable og F-statistik til at teste for samlet signifikans. Normalitets antagelsen testes i dette projekt hver gang en t- eller F-statistik ønskes anvendt til at vurdere signifikans af estimater. Dette gøres ved brug af både en Shapiro-Wilk normalitetstest og en Jarque Bera normalitetstest. Den anvendte model overholder denne antagelse og resultatet af normalitetstesten af fejlleddet kan ses i appendiks tabel A.5.

Ud over de seks antagelser er modellernes estimater også undersøgt for værende tidskonsistente ved brug af CUSUM-test, som beregner kumulative summer af residualer eller rekursive residualer. Testen blev udført ved at akkumulere residualer fra de tidligste observationer og fortsætte denne proces, indtil alle observationer var inkluderet. Over hele perioden fremgår et bånd på $\pm$ 2 standardafvigelser. Det er hermed muligt at teste, om modellens estimater er tidsinkonsistente eller om der opleves strukturelle ændringer i den undersøgte tidsperiode. Resultater af denne test kan ses i appendiks figur A.3. Her udledes

det fra den OLS-baserede CUSUM-test, at der ikke er tidsinkonsistente estimater i modellen, men fra recursive CUSUM-testen går modellen kortvarigt udover båndene, hvorfor der her kunne være et potentielt strukturelt brud i estimaterne på det specifikke tidspunkt. Siden linjen ikke viser et konsistent mønster i at være udenfor båndene er der dog ikke tale om et entydigt brud på tidskonsistensen. Samlet vurderes de to CUSUM-tests at vise, at estimaterne er tidskonsistente.

3.6 Difference-in-Differences

I en Difference-in-Differences (DiD) analyse bliver en kausal sammenhæng mellem treatment og en outcome variabel forsøgt undersøgt. Treatment kan være et specifikt drastisk udfald i en bestemt variabel eller en anvendt politisk reform. I sådan en analyse er der en treatmentgruppe og en kontrolgruppe. Treatmentgruppen er den gruppe, der bliver påført ændringen - eller udfaldet. Kontrolgruppen er derimod en gruppe, der ligner treatmentgruppen, men som ikke har oplevet behandlingen. Hvis man har fundet en god kontrolgruppe vil man således kunne udlede effekten på outcome variablen af behandlingen på treatmentgruppen (de behandlede). Et eksempel herpå kunne være at undersøge effekten af en stigning i prisen på privatskoler i én region, mens en nærliggende region ikke har oplevet nogen ændring. For at udlede den faktiske effekt af behandlingen skal man vide, hvordan treatmentgruppen havde reageret, hvis den ikke havde modtaget behandlingen. Da det ikke er muligt for treatmentgruppen både at få behandlingen og ikke at få behandlingen kan en counterfactual opstilles. Den måler, hvad outcome variablen (i dette eksempel: prisen på privatskolerne i regionen) i treatmentgruppen havde været, hvis der ingen behandling havde fundet sted (Gregory, 2022). Denne counterfactual illustreres lettest ved en graf, hvor Y angiver outcome variablen, mens den stiplede vertikale linje angiver, hvornår behandlingen indtraf.



Figur 3: Counterfactual i en DiD-analyse.

Illustrationen i figur 3 viser, at Y for counterfactual er 8 i år 2018, mens Y for treatmentgruppen i 2018 er 9. Således vil DiD-estimatoren, Δ^{DiD}, være 1. DiD-ligningen gives ved:

$$\Delta^{DiD} = \underbrace{E[Y|G=1,T=1] - E[Y|G=1,T=0]}_{\text{Tidsforskellen i treatmentgruppen}} - \underbrace{E[Y|G=0,T=1] - E[Y|G=0,T=0]}_{\text{Tidsforskellen i kontrolgruppen}}$$

Hvor Y er outcome variablen, G angiver hvilken gruppe der er tale om ($G = 1$ angiver således treatmentgruppen, mens $G = 0$ angiver kontrolgruppen) og T angiver tiden (her angiver $T = 1$ tiden efter behandlingen indtraf, mens $T = 0$ angiver tiden før treatment indtraf).

Første led af ligningen angiver, som illustreret, tidsforskellen i treatmentgruppen, mens sidste led angiver tidsforskellen i kontrolgruppen.

For at kunne udlede en valid counterfactual er det vigtigt, at antagelsen om parallelle trends er overholdt. Denne antager, at treatmentgruppen og kontrolgruppen har ens udvikling i trend for outcome variablen i perioden op til at behandlingen indtraf. Dermed skal følgende ligning være overholdt for at kunne anvende den udledte counterfactual:

$$E[\varepsilon|G=1,T=1] - E[\varepsilon|G=1,T=0] = E[\varepsilon|G=0,T=1] - E[\varepsilon|G=0,T=0]$$

Det skal således gælde, at hvis behandlingen ikke havde fundet sted, så skal udviklingen i forskellene mellem treatment- og kontrolgruppen som ikke er observerbare være ens over

tid (Columbia University, 2013).

Der kan i en DiD-analyse være fokus på at opnå forskellige estimater. Et fokus kan være at forsøge at estimere Average Treatment Effect (ATE), som i modsætningen til Average Treatment Effect on the Treated (ATET) er den samlede effekt, der ses fra treatment fordelt ud på populationen. I ATE er det altså ikke blot effekten på den undersøgte treatmentgruppe, der undersøges, som det ellers er tilfældet i ATET. Den typiske tilgang er dog at forsøge at estimere ATET, som også er den metode, der anvendes i dette projekt. Til udregningen af dette er en metode at anvende en regression af følgende format (Alves, 2022):

$$Y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Treated_i + \beta_3 Treated_i * Post_t + \varepsilon_{it}$$

Her er $Post_t$ en binær tidsvariabel, der tager værdien 1, hvis der er tale om post-treatment og 0, hvis der er tale om præ-treatment. Samtidig er $Treated_i$ en binær gruppevariabel, der tager værdien 1, hvis der er tale om treatmentgruppen og 0, hvis der er tale om kontrolgruppen. Dermed fanger estimatet til interaktionsleddet mellem disse binære variable ($Treated_i * Post_t$) DiD-effekten, ATET.

3.7 Syntetisk Difference-in-Differences

Det er ikke altid muligt at indsamle en tilhørende kontrolgruppe til treatmentgruppen. I dette projekt forsøges der fx at blive dannet en kontrolgruppe tilhørende Danmark. Her findes der ingen enkelt oplagt kandidat, hvorfor der i stedet forsøges at opstille en syntetisk kontrolgruppe.

3.7.1 Syntetisk kontrolmetode

Synthetic Control Method (SCM) er en statistisk teknik, der bruges indenfor kausalanalyse til at danne en syntetisk kontrolgruppe. Denne syntetiske kontrolgruppe kan dannes ud fra en matching-metode. Her bruges vægte til at finde ligheder mellem treatmentgruppen og mulige donor-lande - altså lignende lande, som ikke har set implementeringen af treatment (i dette projekts tilfælde implementeringen af de afdragsfrie lån), som kan bidrage til den syntetiske kontrolgruppe. Formålet hermed er at danne en gruppe, der minder mest muligt om treatmentgruppen forud for behandlingen (Abadie et al., 2011).

I modsætning til andre matching-metoder anvender SCM både de uafhængige variable i præ-treatment perioden og outcome-variablen til at matche donor-landene med treatmentgruppen (Nguyen, 2023). Således findes de donor-lande, der minder mest om treatmentgruppen. Matching-processen sker for hver periode, så outcome-variablen bliver en tidsserie med udregnede værdier for hver periode. En stor fordel ved SCM er, at den ikke anvender ekstrapolation. Herved skal forstås, at metoden kun anvender den tilgængelige information (data) i sin udregning af den syntetiske kontrolgruppe - ingen forecast eller lignende udføres dermed til at udregne post-treatment perioden for kontrolgruppen, hvilket i sidste ende medfører mere robuste resultater.

Matching og SCM hjælper også imod selection bias, da der anvendes flere passende kontrolgrupper og fordi vægtene af disse udregnes på en objektiv metode.

SCM tager allerede højde for forskelle i kontrolgruppen og treatmentgruppen, så målet med den syntetiske kontrolgruppe bliver i præ-treatment perioden, at forskellen mellem de to grupper bliver så lav som muligt.

Synth-pakken i R er en af flere metoder, der kan anvendes til at implementere SCM i praksis. Her er en mulig anvendelig algoritme Broyden-Fletcher-Goldfarb-Shanno (BFGS) metoden. Denne algoritme minimerer den multivariate funktion, som kovariatene udgør, for derved at finde den kombination af vægte, som minimerer afstanden mellem kontrolgruppen og treatmentgruppen.

3.7.2 Kombinationen af SCM og DiD

Arkhangelsky et al. (2016) undersøgte i 2019 muligheden for at kombinere den ordinære DiD-analyse med SCM. I den traditionelle DiD-analyse analyseres paneldata typisk med mange observationer i både treatment og kontrolgruppen, mens der i SCM typisk er få observationer tilgængeligt. En blanding af disse to metoder kan således gøre op for de manglende observationer i kontrolgruppen.

3.8 Data

Dette projekt anvender primært data på kvartalsbasis, hvilket sker på baggrund af anvendelsen af makroøkonomiske variable, der i de flestes tilfælde kun rapporteres på års- og kvartalsbasis, og fordi der ønskes så mange observationer som muligt indenfor den

analyserede tidshorisont.

Variabel	**Detaljer**	**Kilde**
Boligprisindeks	Enfamilieshuse, hele landet	Abildgren, Kim. A Chart & Data Book on the Monetary and Financial History of Denmark, tabel S030Q
Bruttonationalprodukt	2010-priser, kædede værdier, (mia. kr.)	Danmarks statistik, tabel NKN1
Privatforbrug	2010-priser, kædede værdier, (mia. kr.)	Danmarks statistik, tabel NKN1
Arbejdsløshed	% af arbejdsstyrken	Abildgren, Kim. A Chart & Data Book on the Monetary and Financial History of Denmark, tabel S041Q
Den effektive rente på realkreditobligationer	Den gennemsnitlige lange obligationsrente for realkreditobligationer i danske kroner.	Abildgren, Kim. A Chart & Data Book on the Monetary and Financial History of Denmark, tabel S002Q
Den effektive rente på danske statsobligationer	Den gennemsnitlige lange obligationsrente for statsobligationer i danske kroner.	Abildgren, Kim. A Chart & Data Book on the Monetary and Financial History of Denmark, tabel S001Q
Indlån i pengeinstitutter	Ultimobalance (mio.kr), Husholdninger	Danmarks statistik, tabel DNPIN og Nationalbankens statistikbank, tabel DNINDL1
Boliger N.111	2010-priser, kædede værdier	Danmarks statistik, tabel NKN1
Andel af afdragsfrie realkreditlån	Andel udregnet ud fra ultimobalance (mio.kr), Husholdninger, Ejerboliger og fritidshuse	Nationalbankens statistikbank, tabel DNRUDDKI og DNEJER
Forbrugertillidsindikatoren	Forbrugerforventninger (nettotal)	Danmarks statistik, tabel FORV1

Tabel 1: Det anvendte data samt kilder til at kunne genskabe dataserierne.

Projektets tidshorisont strækker sig fra første kvartal 1994 til fjerde kvartal 2018, da der ikke findes observationer før 1994 med det anvendte datagrundlag og effekter fra coronakrisen ønskes fjernet for ikke at påvirke de analyserede estimater og sammenhænge. Nogle variable såsom opsparing, forbrugertillid og inflation har kun været muligt at finde

på månedsbasis, hvorved en omskrivning til kvartalsbasis har været nødvendigt. I alle tilfælde er gennemsnittet i de tre måneder tilhørende hvert kvartal blevet anvendt som den nye variabel, for på den måde bedst at fange den fulde effekt i kvartalet.

De anvendte tidsserier vil blive beskrevet herunder samt en forklaring af eventuel feature engineering, der er blevet anvendt for at sikre, at alle tidsserier har samme indeks år.

Reale huspriser:

- Denne serie dækker over et nominelt prisindeks for ejendomssalg i hele Danmark for enfamilieshuse oprindeligt med basis år i 2006, men reindekseret til år 2010 for at alle anvendte variable har samme basisår. Dernæst blev serien deflateret med forbrugerprisindekset ligeledes med basisår i år 2010 for at fjerne inflationære ændringer i serien.(Abildgren, 2017)

Bruttonationalproduktet:

- Denne serie dækker over det samlede danske bruttonationalprodukt i millioner kr. i 2010 konstante priser.(Danmarks Statistik, 2023a)

Privatforbrug:

- Denne serie dækker over det samlede danske forbrug i millioner kr. i 2010 konstante priser.(Danmarks Statistik, 2023a)

Arbejdsløshed:

- Denne serie dækker over fuldtids nettoledige i Danmark som % af den samlede arbejdsstyrke.(Abildgren, 2017)

Den effektive rente på realkreditobligationer:

- Denne serie dækker over den effektive rente på danske realkreditobligationer med 30 års løbetid.(Abildgren, 2017)

Den effektive rente på danske statsobligationer:

- Denne serie dækker over den effektive rente på danske statobligationer med 10 års løbetid. (Abildgren, 2017)

Opsparing:

- Denne serie dækker over indlån fra danske husholdninger i danske pengeinstitutter i alt i millioner kr. Serien er lavet fra månedsbasis til kvartalsbasis og derefter deflateret med forbrugerprisindekset ligeledes med basisår i år 2010 for at fjerne inflationære ændringer i serien.(Danmarks Statistik, 2023b) (Nationalbankens Statistikbank, 2013b)

Boligmængde:

- Denne serie dækker over en proxy for boligmængden, hvor N.111 fra Danmarks nationalregnskab er blevet anvendt som værdien af den samlede mængde boliger i Danmark. Denne serie divideres med det allerede klargjorte reale boligprisindeks for at skabe en proxy for boligmængden i Danmark. (Abildgren, 2017)

Andel af afdragsfrie realkreditlån:

- Denne serie dækker over den samlede værdi af afdragsfrie realkreditlån udstedt til husholdninger i Danmark divideret med den samlede værdi af alle typer realkreditlån udstedt til husholdninger i Danmark. Serien vil derfor være 0, hvis ingen danske husholdninger anvender afdragsfrie lån, som det er tilfældet frem til fjerde kvartal 2003, hvor det blev muligt at optage afdragsfrie realkreditlån. Derefter vil serien forblive mellem 0 og 1 afhængig af populariteten på låntypen.(Nationalbankens Statistikbank, 2023) (Nationalbankens Statistikbank, 2013a)

Forbrugerforventninger:

- Denne serie dækker over danske forbrugeres forventninger til deres egen families samt lands økonomiske situation. Denne serie er lavet fra månedsbasis til kvartalsbasis.(Danmarks Statistik, 2023c)

Efterfølgende er den naturlige logaritme blevet anvendt på alle dataserier. Dette er blevet gjort på reale huspriser, bruttonationalproduktet, forbrug, arbejdsløshed, den effektive rente på realkreditobligationer og statsobligationer, opsparing og boligmængde. Det har ikke været muligt at lave transformationen for andel af afdragsfrie realkreditlån dataserien samt forbrugerforventninger, da disse begge indeholder observationer med negative

værdier eller 0, hvilket gør det matematisk umuligt.

Til analysen med et DiD-framework er der grundet metodemæssige restriktioner anvendt årligt data i stedet for kvartalsvist data. Selve opbygningen af data er også forskellig, da den økonometriske analyse anvender tidsseriedata, mens DiD-analysen anvender panel-data.

De anvendte variable i paneldataet fremgår af tabel 2 og løber over perioden 1996-2018.

Variabel	**Detaljer**	**Kilde**
BNP	Bruttonationalprodukt til markedspriser. Prisindeks med 2010 = basisår.	Eurostat Kode: nama_10_gdp unit: PD10_EUR na_item: B1GQ
Arbejdsløshed	Arbejdsløshedsraten i % af populationen	Eurostat Kode: une_rt_a_h unit: PC_ACT age: Y15-74 sex: Total
Rentesats	EMU konvergenskriterieserien for obligationsafkast	Eurostat Kode: irt_lt_mcby_a int_rt: MCBY
Population	Population i tusinde personer.	Eurostat Kode: nama_10_pe unit: THS_PER na_item: POP_NC
Boligpriser	Boligprisindeks i reale priser. CPI-deflateret.	The Bank of International Settlements (BIS, 2023)

Tabel 2: Det anvendte data til SCM og fixed effects modellerne samt kilder til at kunne genskabe dataserierne.

Det har været nødvendigt at indsamle variable for mange lande, hvorfor Eurostat er blevet anvendt som primær datakilde - dog er husprisindekset fundet gennem *The Bank of International Settlements*. Dette øger sammenligneligheden af hver enkelt variabel, da samme metode er anvendt til at måle de enkelte variable.

Efter at have redegjort for projektets metoder og data, er det nu tid til at udvælge de relevante variabler til brug i analysen. Dette vil blive udført i det følgende afsnit.

4 Deskriptiv analyse

I sidste afsnit blev den økonometriske projektmetode opstillet, mens relevant data blev præsenteret. I dette afsnit vil der blive sorteret i det fundne data vha. dimensionality reduction metoder.

4.1 Udvælgelse af variable

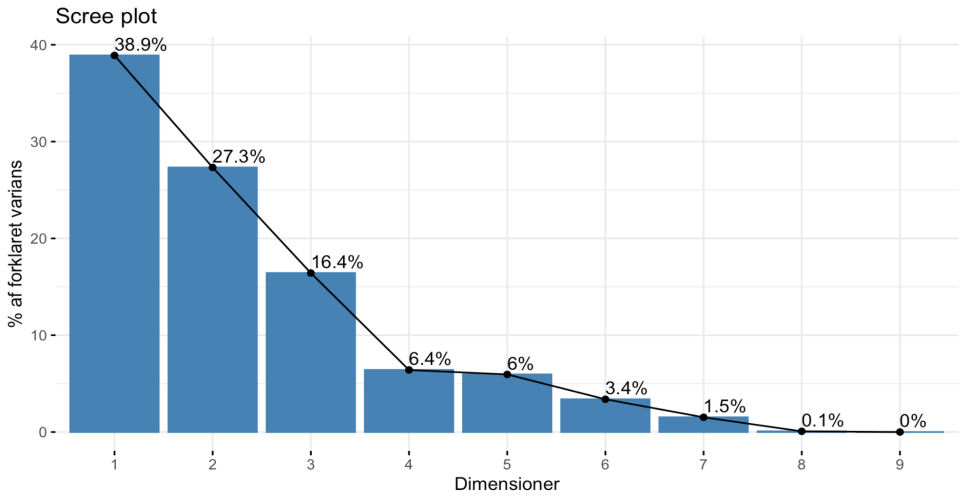
4.1.1 PCA

Der er i projektet blevet udført en PCA. Denne er blevet anvendt med henblik på at fjerne nogle af variablene fra analysen og dermed skabe en lavere grad af dimensionality samtidig med at mest muligt information beholdes i datasættet. Analysen blev foretaget på variablene: BNP, Forbrug, Opsparing, Arbejdsløshed, Tillid, boligmængde, 10-årige statsobligationsrenter, 30-årige realkreditobligationsrenter og Andel (af samlet lån udgjort af de afdragsfrie lån). Dataet blev først normaliseret, hvorefter en korrelationstest blev kørt. Resultaterne heraf ses af figur A.2 i appendiks.

Herfra udledes det, at variablene log_forbrug og log_bnp ser ud til at være korrelerede, hvorfor det allerede her noteres, at kun en af disse variable vil blive anvendt i ARDL-Bounds testen. Der kunne på forhånd være en mistanke om, at rentevariablene ville være stærkt korrelerede. Mistanken blev bekræftet gennem korrelationstesten, hvorfor en af disse heller ikke vil blive anvendt i den endelige model.

Efter korrelationerne blev defineret, blev først princomp() funktionen og dernæst prcomp() funktionen, som er en nyere og mere kraftfuld PCA metode i R, anvendt for at danne fundamentet for PCA. Af denne grund blev prcomp() metoden foretrukket, hvorfor det er denne metode, der vil blive gennemgået i dette afsnit.

Mængden af dimensioner og disses tilhørende varians kunne visualiseres i et scree-plot, som kan ses i figur 4.



Figur 4: Scree-plot over variansen for de forskelle PC's.

Det blev tidligere nævnt, at der er tre metoder til at udvælge antallet af dimensioner i en PCA: Scree-test, egenværdier eller et fastsat mål for kumuleret varians. Fra scree-testen ses "albuen" tydeligt ved 4. dimension. Således vælges de antal dimensioner til venstre for albuen ud fra scree-testen. Yderligere kan valget bakkes op ved tabel 3, hvor egenværdierne for hver dimension vises. derudover kan den kumulative varians tilhørende disse dimensioner udledes fra tabellen.

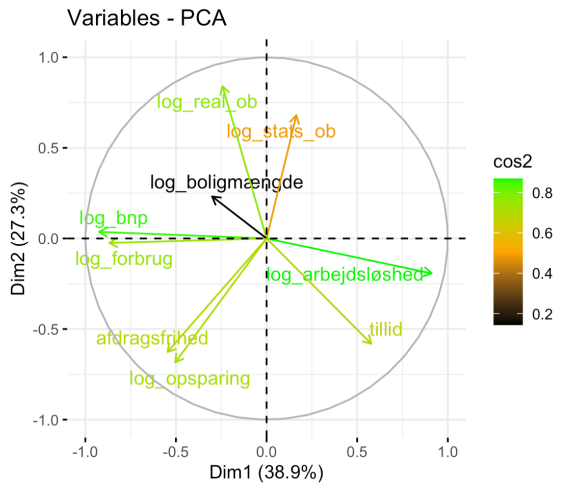
	Egenværdi	**Varians.procent**	**Kumulativ.varians.procent**
Dim.1	3.500	38.900	38.900
Dim.2	2.460	27.300	66.200
Dim.3	1.480	16.400	82.700
Dim.4	0.577	6.410	89.100
Dim.5	0.536	5.950	95.000
Dim.6	0.305	3.390	98.400
Dim.7	0.137	1.520	99.900
Dim.8	0.007	0.073	100
Dim.9	0	0	100

Tabel 3: Egenværdier for dimensionerne.

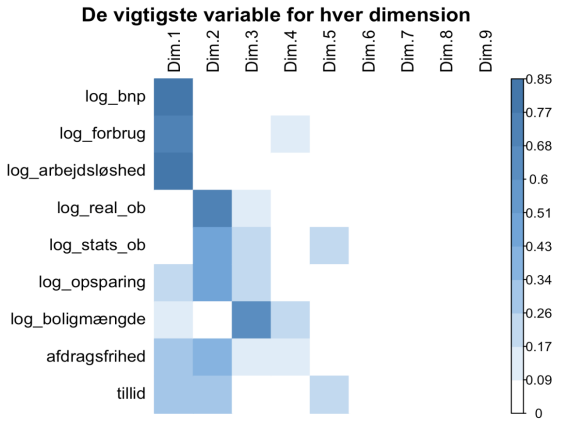
Det ses fra tabellen, at dimension 1-3 alle er større end 1, mens dimension 4-9 alle er under 1 i egenværdi. Egenværditesten foreslår dermed, at datasættet skal fremvises i 3 dimensioner. Dette antal dimensioner er således det antal, der anvendes i analysen. Ved dette antal dimensioner er 82.7% af variansen forklaret.

Næste trin var at finde de vigtigste variable indenfor de 3 dimensioner. Til dette

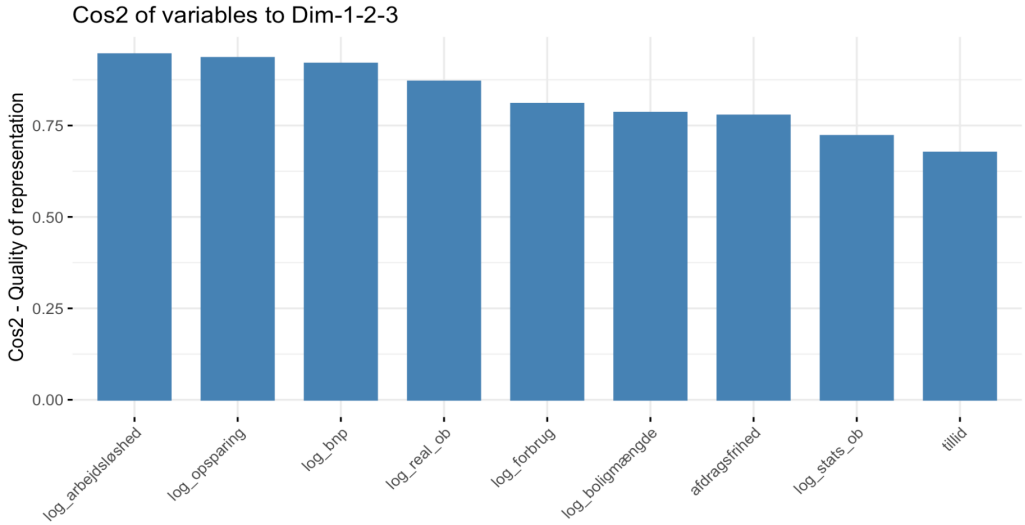
formål kunne illustrationerne i figur 5 anvendes.



(a) 2-dimensionelt plot af vigtigheden af de forskellige variable.



(b) Illustration af de vigtigste variable fordelt ud på dimensionerne jf. PCA.



(c) Illustration af de samlet vigtigste variable jf. PCA.

Figur 5: Samlet plot af 3 forskellige PCA variabel plots.

Fra figur 5a ses de to vigtigste dimensioner illustreret, hvor hver variabel og dens tilhørende egenvektor ligeledes er illustreret. Her er længden af egenvektorerne givet ved egenværdierne. Derudover angiver farven på variablen dets kvadrerede cosinuskorrelation med PC's. Sort er en lav korrelation, mens grøn svarer til en stærk korrelation. For at fortolke figuren skal man betragte x-aksen som repræsentationen for sammenhængen mellem variablen og den første dimension, mens længden af egenvektorerne op ad y-aksen angiver sammenhængen mellem variablene og den anden dimension. Således vurderes det, at log_boligmængde er den mindst vigtige variabel til at forklare de to første dimensioner,

da den er mindst korreleret med de opstillede PC's og er den korteste vektor.

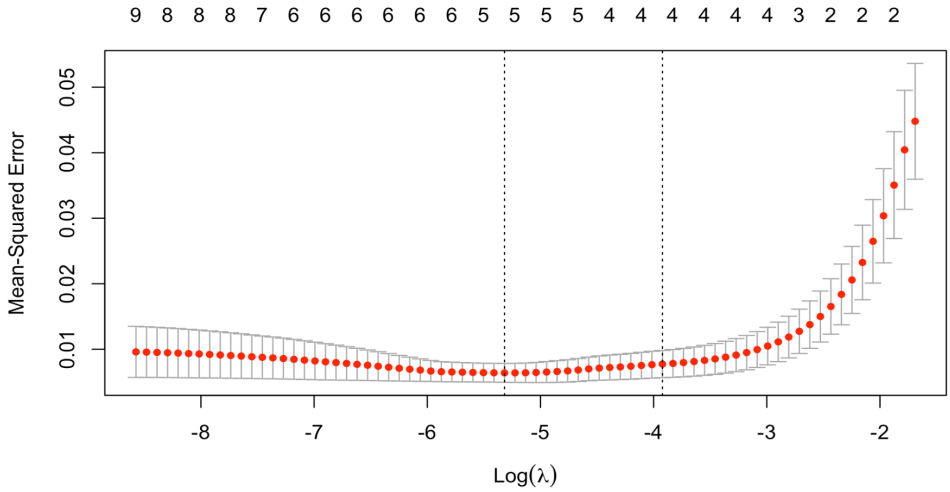
En oversigt over alle anvendte variables vigtighed for hver dimension fremgår af figur 5b. Jo mørkere farven er, jo større betydning har variablen for den pågældende dimension. Dermed er log_bnp og log_arbejdsløshed de vigtigste variable i dimension 1, mens log_real_ob, log_stats_ob og log_opsparing er de vigtigste variable til at forklare dimension 2. Slutteligt er log_boligmængde den vigtigste variabel til at forklare dimension 3. Selvom denne figur er god til at forklare vigtigheden af variablen for den givne dimension, så siger den ikke meget om de samlet vigtigste variable på tværs af de tre første dimensioner.

Til en samlet vurdering kan figur 5c anvendes; denne viser netop dette, igen beregnet ud fra variablenes kvadrerede cosinuskorrelation til dimension 1-3. Dermed ender de mindst vigtige variable med at være tillid, log_stats_ob, afdragsfrihed og log_boligmængde. Med undtagelse af afdragsfrihed (da denne er analysens omdrejningspunkt) vurderes det således, at disse variable kan udelades fra analysen uden større tab af forklaringskraft. Samtidig blev det tidligere fremhævet, at forbrug og BNP har en stærk korrelation til hinanden, hvorfor en af disse variable også droppes. Samlet set fra figur 5 vurderes forbrug at være den mindst vigtige variabel af disse, hvorfor denne fravælges.

4.1.2 LASSO-regression

Som en yderligere dimensionality reduction metode har projektet anvendt en LASSO-regression som rygdækning til PCA metoden. Her udføres først en krydsvaliderings test for at bestemme den optimale λ-værdi. Træningen af modellen fremgår af figur 6.

Denne træning har endt ud i, at den optimale λ-værdi for regressionen er 0.0049. Denne værdi bruges dermed til at udregne straffen til variabelkoefficienterne. PCA fremgav tillid, log_stats_ob, andel og log_boligmængde som de mindst vigtige variable. Af tabel 4 fremgår de mindst vigtige variable jf. LASSO-regressionen.



Figur 6: Krydsvaliderings resultaterne.

De mindst vigtige variable er de variable med nulværdier, det vil sige log_boligmængde, log_opsparing, log_real_ob og log_forbrug. Dermed kan det udledes på baggrund af de to dimensionality reduction analyser, at de eneste variable, der er blevet valgt på baggrund af begge analyser er afdragsfrihed, arbejdsløshed og BNP. Således bliver risikoen for multikollinearitet mindsket, mens også kun de vigtigste variable nu er fastholdt.

Variable	Koefficienter	Udvalgte
log_arbejdsløshed	-0.108	Yes
log_bnp	0.084	Yes
Afdragsfrihed	0.040	Yes
log_stats_ob	0.030	Yes
Tillid	0.029	Yes
log_forbrug	0.000	No
log_real_ob	0.000	No
log_opsparing	0.000	No
log_boligmængde	0.000	No

Tabel 4: Variabel selektion ved brug af LASSO regressionen.

4.2 Udvælgelse af lande til Syntetisk DiD

Udvælgelse af lande til at indgå i den syntetiske kontrolgruppe sker på baggrund af landenes anvendelse af afdragsfrihed på deres respektive realkreditmarkeder. Det er værd at nævne, at Danmarks realkreditmodel med specialbanker (realkreditinstitutter), der kun må yde lån mod sikkerhed i fast ejendom finansieret med realkreditobligationer er ret unik, også i internationalt perspektiv, hvorved det kan være svært at sammenligne med

andre lande en til en (Realkredit Rådet, 2015). Derfor vil følgende afsnit ikke som udgangspunkt tage hensyn til opbygningen af diverse europæiske landes realkreditmodeller, men i større grad i hvilken grad afdragsfrie produkter bliver anvendt af husholdningerne til at finansiere boliglån.

De lande der indgår i udvælgelsesprocessen er alle europæiske, da det er vurderet, at disse lande i større grad minder om Danmark end lande fra resten af verden. Derudover har det anvendte datagrundlag fra perioden 1996 til 2018 også sat en begrænsning for fri anvendelse af samtlige lande i Europa, hvilket har skabt en naturlig selektion. De lande der vil indgå i udvælgelsesprocessen er Belgien, Finland, Holland, Italien, Irland, Norge, Spanien, Sverige og Tyskland. Disse lande vil nu kort blive gennemgået enkeltvis for at få et indblik i anvendelse af afdragsfrihed på deres realkreditmarked.

Belgien: I Belgien er afdragsfrie lån ikke forbudt ved lov og forekommer hovedsageligt i to kategorier. Mellemfinanseringslån, der hovedsageligt bruges imellem køb af ny bolig og salg af nuværende og et stående lån, der er et afdragsfritlån, hvor der kun betales renter og hovedstolen tilbagebetales ved lånets udløb. Disse to låntyper udgjorde i 2016 11% af den samlede mængde af udstående boliglån i Belgien, hvor 7% kommer fra mellemfinanseringslån, der kan siges at være lån med kort løbetid på maks 2 år, hvorved kun en meget lille del af boliglånene i Belgien finansieres ved afdragsfrihed. (European Mortgage Federation, 2019)

Finland: I Finland er afdragsfrie lån ikke en populær lånetype, hvor det hovedsageligt bliver brugt til "buy-to-let" lån, der er et banklån, hvor udgangspunktet for lånet er at leje den købte ejendom ud og dermed anvende den som investeringsgrundlag og ikke som beboelse. Data for hvor stor en mængde af disse lån, der anvendes i Finland er først begyndt at blive indsamlet i år 2021, og har lige siden lagt på omkring 8% af den samlede mængde udlån. (Finlands Bank, 2023)

Italien: I Italien er afdragsfrie realkreditlån ikke udbredt og slet ikke anvendt over en 30-årig låneperiode. Det er muligt at få perioder, hvor der ikke afdrages men det er ikke muligt at estimere antallet af lån, der benytter sig af denne praksis. (European Mortgage Federation, 2019) (Frank van Hoenselaar et al., 2021)

Irland Siden begyndelsen af 2013 har mindre end 1% af de nye realkreditlån i Irland været afdragsfrie, mens resten er fuldt ud amortiseret. Udlån med afdragsfrihed toppede i 2006 i mængde og i 2007 som andel af samlede udlån. Den største anvendelse blev set hos låntagere, der købte til boligudlejning eller anvendte lånet som mellemfinanseringslån. (Department of Housing og Government, 2017)

Norge: I Norge anvendes afdragsfrihed primært på en låntype kaldet rammelån, hvor der sættes en ramme på 60% af boligens værdi som er mulig at låne i. Lånet er i princippet afdragsfrit, da låntager selv bestemmer, hvor store afdrag hvis overhovedet, der betales hver periode, hvilket kan gøre det vanskeligt at vurdere den reelle mængde, der anvender dem som afdragsfrie lån. Låntypen toppede i 2012 på omkring 27% af den samlede mængde udlån. (Statistics Norway, 2023)

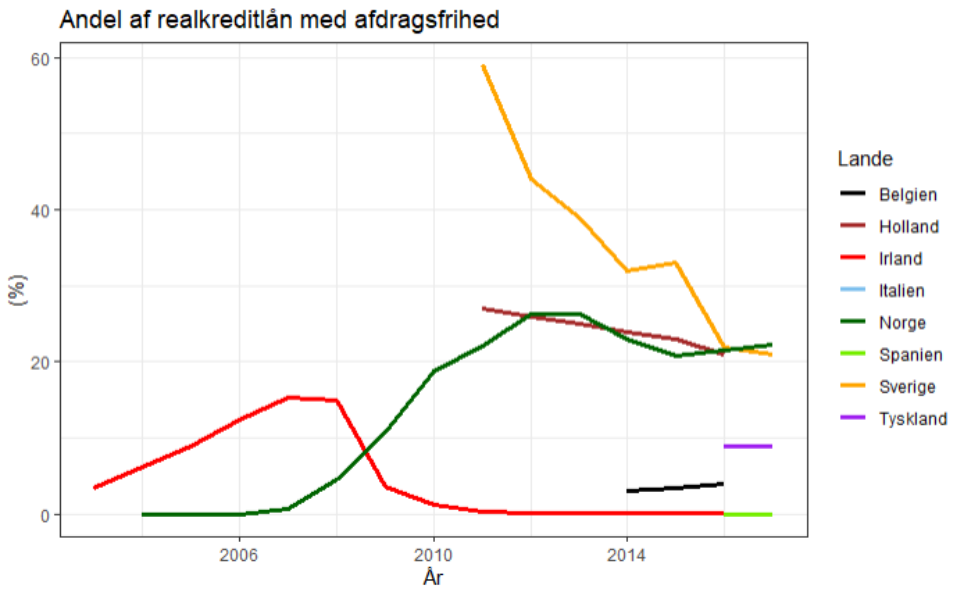
Spanien: I Spanien er det ikke muligt at få afdragsfrie realkreditlån på nuværende tidspunkt. Det har heller ikke været muligt at finde yderligere data, der understøtter, at det nogensinde skulle have været en alment brugt lånetype, hvilket kan skyldes datamangel eller, at det reelt set aldrig har været udbredt i landet. (Lionsgate Capital, 2023)

Sverige: I Sverige er afdragsfrie realkreditlån en hyppigt anvendt låntype. Dog har anvendelsen svinget meget fra næsten ingen anvendelse i 1990'erne til at være den mest populære låntype i landet siden 2011. Dog kom der ny lovgivning i 2015, der har sat restriktioner på optagelse af lånene, hvilket har gjort, at andelen af den samlede mængde udlån, der går til afdragsfrie lån er faldet siden sin top i 2011 og i 2017 var nede på 21%. (European Mortgage Federation, 2019) (Finansinspektionen, 2017)

Tyskland: I Tyskland er afdragsfrie lån ikke en populær lånetype, hvor det hovedsageligt bliver brugt til "buy-to-let" lån, der er et banklån, hvor udgangspunktet for lånet er at leje den købte ejendom ud og dermed anvende den som investeringsgrundlag og ikke som beboelse. Andelen af disse lån lå i 2016 på 9%, dog har det ikke været muligt at finde yderligere data for udviklingen. (Frank van Hoenselaar et al., 2021)

Af figur 7 kan udviklingen i anvendelsen af afdragsfrie realkreditlån ses for ovenstående lande i det omfang, at datapunkter har været tilgængelige. Datapunkter fra Sverige er andelen af nyoptagede realkreditlån, der er afdragsfrie imodsætning til den samlede

mængde af udstående lån.



Figur 7: Anvendelse af afdragsfrie realkreditlån i udvalgte lande. obs. Italien og Spanien ligger oveni hinanden, hvilket medfører, at man ikke kan se linjen for Italien.

Baseret på den ovenstående gennemgang af landenes brug af afdragsfrie realkreditlån, er det besluttet at ekskludere de tre lande, der har den største anvendelse af denne låntype fra yderligere brug til dannelse af den syntetiske kontrolgruppe. Disse tre lande inkluderer Holland, Norge og Sverige.

Dermed konkluderes variabel- og landeudvægelsen. Dette leder i næste afsnit først til en opstilling af en DiD-analyse baseret på en syntetisk kontrolgruppe og dernæst til, at det nye datasæt anvendes til en ARDL bounds test.

5 Analyse

Følgende sektion vil præsentere projektets resultater og forsøge at give et svar på i hvilket omfang introduktionen af afdragsfrie realkreditlån har haft en indvirkning på udviklingen i de danske reale huspriser.

5.1 Syntetisk Difference-in-Differences

En metode til at udregne effekten af implementeringen af de afdragsfrie lån i Danmark i 2003 på huspriser er at anvende en DiD-approach. Da der ikke er nogle åbenlys kontrolgruppe til Danmark, vil der i stedet for en DiD-analyse blive udført en Syntetisk Difference-in-Differences (SDiD) analyse. Hertil skal først estimeres en syntetisk kontrolgruppe.

5.1.1 SCM

Til at bidrage til udvælgelsen af lande til en syntetisk kontrolgruppe er sektion 4.2 opstillet. Da den syntetiske kontrolgruppe gerne skal have haft så lav anvendelse af afdragsfrie lån som muligt, udgør den sektion det historiske fundament for landenes brug heraf.

Til dannelsen af den syntetiske kontrolgruppe er Synth() funktionen blevet anvendt i R. Denne kræver paneldata, en outcome variabel samt nogle udvalgte uafhængige variable. Dette anvendes til at sammenligne, hvilke lande der minder mest om Danmark med hensyn til de udvalgte variable. Hertil danner modellen vægte, der i sidste ende udregner hvor meget hvert lands husprisindeks skal vægtes i udregningen af den syntetiske kontrolgruppes husprisindeks. Modellen er blevet trænet med de i sektion 4 udvalgte variable og på tidshorisonten 1996-2003.

Vægt	**Land**
0.058	Belgien
0.044	Finland
0.009	Tyskland
0.082	Irland
0.422	Italien
0.385	Spanien

Tabel 5: Landevægte i den syntetiske kontrolgruppe.

Landevægtene anvendt til den syntetiske kontrolgruppe fremgår af tabel 5.

Det kan udledes, at den syntetiske kontrolgruppe har vægtet Italien højest med en vægt på 0.422. Således udgøres 42.2% af kontrolgruppen af husprisudviklingen i Italien. Samtidig ses det, at Spanien er vægtet næsthøjest, 38.5%, mens Irland, Belgien og Finland er vægtet med hhv. 8.2%, 5.8% og 4.4%. Således anvendes Tyskland kun i et meget begrænset omfang til fremstillingen af den syntetiske kontrolgruppe.

Tabel 6 angiver også vægte til kontrolgruppen - dog for de anvendte variable.

Variabel	**Vægt**
BNP-indeks	0.913
Arbejdsløshedsrate	0
Rentesats	0.079
Population (i 1000)	0.007

Tabel 6: Vægte for variable i den syntetiske kontrolgruppe.

Det ses tydeligt, at BNP-indekset er den dominerende variabel til dannelsen af den syntetiske kontrolgruppe. Variablen vægtes med 91.3%, mens rentesatsen vægtes med 7.9%. Det er således primært disse variable, som SCM anvender til at skabe en syntetisk kontrolgruppe for Danmark. Arbejdsløshedsraten anvendes slet ikke, og populationen anvendes kun minimalt og har derfor en negligérbar effekt på dannelsen af kontrolgruppen.

Efter vægtene er blevet udregnet, kan den syntetiske kontrolgruppe nu dannes. Tabel 7 viser i denne forbindelse sammenhængen mellem de uafhængige variable for treatmentgruppen og den syntetiske kontrolgruppe.

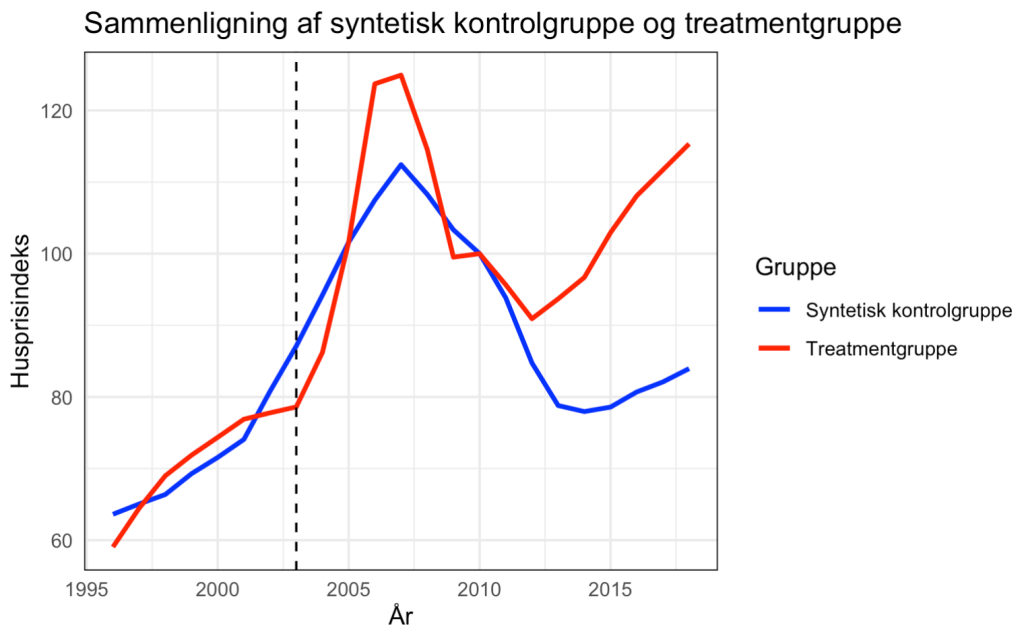
Variabel	**Treatmentgruppe**	**Syntetisk kontrolgruppe**
BNP	78.243	78.270
Arbejdsløshedsrate	5.083	11.860
Rentesats	5.313	5.294
Population (i 1000)	5330.000	41496.888

Tabel 7: Sammenligning af treatmentgruppen og den syntetiske kontrolgruppe.

Sammenligningen fremhæver lighederne mellem de to grupper for de fire anvendte forklarende variable i præ-treatment perioden. Dermed kan det udledes, at gennemsnittet for BNP-variablen i præ-treatment perioden har et gennemsnit på 78.243, mens den synteti-

ske kontrolgruppes gennemsnitlige BNP-indeks i præ-treatment perioden ligger på 78.270. Således er denne variabel meget sammenlignelig for de to grupper. Det samme kan siges om rentesatsen, som ligger på hhv. 5.313 for treatmentgruppen og 5.294 for den syntetiske kontrolgruppe. Dog er den syntetiske kontrolgruppes arbejdsløshedsrate over dobbelt så høj som for treatmentgruppen, mens populationen i den syntetiske kontrolgruppe er næsten 8 gange så stor som den i treatmentgruppen. Dette skyldes til dels, at der i den syntetiske kontrolgruppe hovedsageligt er lagt vægt på Italien og Spanien.

Den syntetiske kontrolgruppe indeholder en estimeret tidsserie for outcome variablen, husprisindekset, fra 1996-2018. Denne visualiseres i figur 8, hvor udviklingen i husprisindekset for både Danmark (treatmentgruppen) og den syntetiske kontrolgruppe kan udledes.

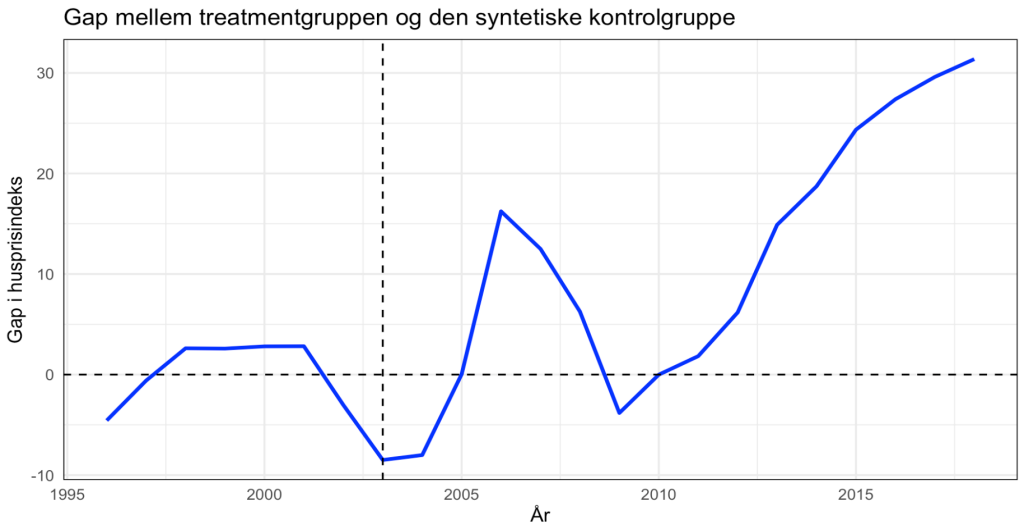


Figur 8: Sammenligning af treatmentgruppe og den syntetiske kontrolgruppes udvikling i husprisindeks.

Vigtigt i denne sammenhæng er det, at præ-treatment værdierne for kontrolgruppen nogenlunde afspejler værdierne for treatmentgruppen. Umiddelbart følges de to tidsserier indtil 2003, hvor treatment fandt sted. Derefter skabes større udsving i tidsserierne, hvor treatmentgruppens husprisindeks vokser hurtigere.

Også figur 9 kan anvendes til at følge udviklingen. Denne viser forskellen mellem de to grupper. Forskellen er udregnet ved $hp_{treatment} - hp_{syntetisk}$, så positive værdier viser

årstal, hvor treatmentgruppens husprisindeks er større end kontrolgruppens husprisindeks.



Figur 9: Forskellen mellem treatmentgruppen (Danmark) og den syntetiske kontrolgruppe.

Her understøttes det, at det især er efter 2003, at Danmarks husprisindeks vokser ift. den syntetisk opstillede kontrolgruppe. SCM er således færdiggjort, og den syntetiske kontrolgruppe er opstillet. Dermed kan denne anvendes til at udføre den SDiD-analysen.

5.1.2 SDiD

Til udførelsen af SDiD-analysen blev der opstillet nye binære variable for tid (1 hvis årstallet er efter 2003 og ellers 0) og gruppe (1 hvis der er tale om treatmentgruppen - Danmark - og 0, hvis der er tale om den syntetiske kontrolgruppe). Disse blev brugt til at danne variablen did (tid*gruppe), som er DiD-estimatoren. Regressionen $hp = gruppe + tid + did$ blev kørt i en LM model. Resultaterne heraf fremgår af tabel 8.

Koefficienter	**Estimater**	**Std. Fejl**	**t værdi**	**Pr(>|t|)**
(Intercept)	72.2324	3.7841	19.089	< 2e-16 ***
gruppe	-0.7334	5.3515	-0.137	0.8917
tid	20.3017	4.6857	4.333	8.98e-05 ***
did	12.5700	6.6266	1.897	0.0647 .

Signifikanskoder:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tabel 8: Estimater for DiD-analysen.

Tiden i sig selv er meget signifikant, hvorfor der samlet for de to grupper kan konkluderes et skifte i husprisindeks fra perioden før treatment til efter treatment indtraf. Den vigtige variabel er dog did variablen. Estimatet herfor er positivt med en værdi på 12.5700, hvilket betyder, at husprisindekset alt andet lige skulle være steget med 12.57 indekspoint som følge af implementeringen af afdragsfrie lån i Danmark tilbage i 2003. P-værdien for undersøgelsen ligger på 0.0647. Hertil skal nævnes, at der typisk anvendes et signifikansniveau på 5%, hvorfor den statistiske signifikans af estimatet er underlagt en vis usikkerhed.

Udover ovenstående DiD, blev yderligere en model opstillet - denne gang vha. plm(), som er en panel lineær model - med det formål at implementere BNP-indekset, arbejdsløshedsraten og rentesatsen som kontrolvariable. Dette ligger til grund for den model, som fremgår af tabel 9. DiD-modellen med kontrolvariable tager højde for de ekstra tilføjede variable foruden dem i standard DiD-analysen, der blev opstillet i tabel 8. Dermed forsøges disse ekstra variable implementeret for potentielt at mindske omitted variable bias.

Koefficienter	**Estimater**	**Std. Fejl**	**t værdi**	**Pr(>|t|)**
tid	10.62067	5.23428	2.0291	0.049318 *
did	18.00357	4.25156	4.2346	0.000135 ***
arbejdsløshed	-3.73138	0.43334	-8.6107	1.479e-10 ***
rentesats	2.97710	1.09583	2.7167	0.009779 **
BNP-indeks	0.74649	0.26585	2.8079	0.007749 **
Signifikanskoder:				
0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' ' 1

Tabel 9: Estimater for DiD-analysen med kontrolvariable.

Der ses en tydelig forbedring af signifikansniveuaerne for de undersøgte variable. Både DiD-estimatoren og arbejdsløshed er signifikant ved et signifikansniveau på 0.1%. Her er estimatet for did 18.00, hvilket antyder, at implementeringen af afdragsfrie lån gennemsnitligt over post-treatment perioden har medført en stigning i husprisindekset på 18.00 indekspoint for treatmentgruppen. Samtidig ses estimaterne for kontrolvariablene, der er blevet introduceret i modellen. Arbejdsløshedsestimatet tyder på en negativ sammenhæng mellem arbejdsløshed og husprisindekset, hvor en 1-procentpoint stigning i arbejdsløshed vil medføre et fald i husprisindekset på 3.73 indekspoint. Derudover påstår modellen, at

en stigning i rentesatsen på et procentpoint vil medføre en stigning i husprisindekset på 2.98. Slutteligt vil en stigning i BNP-indekset på ét indekspoint medføre en stigning i husprisindekset på 0.75 indekspoint.

For at tjekke resultaterne heraf er et robusthedstjek blevet udført. Tabellen hertil fremgår i appendiks D.1. Denne model er opbygget på baggrund af årene 1996-2010 i stedet for 1996-2018. Dette er for at teste, om der sker et skifte i signifikans og estimat ved ændringer i modelspecifikationen - i dette tilfælde i form af tidshorisonten. Robusthedstjekket ændrer ikke på estimaternes fortegn, hvor did-estimatet er 13.79, arbejdsløshedsestimatet er -2.64, renteestimatet er 1.31 og BNP-estimatet er 0.96. Der ses stadig signifikante estimater, udover rentesatsen, som ikke længere er signifikant ved et signifikansniveau på 10%. Dog er analysens omdrejningspunkt, DiD-estimatoren, signifikant på et 1% signifikansniveau.

5.2 ARDL Bounds test

Sektion 4 fungerede som variabel udvælgelse, hvor variablene BNP, arbejdsløshed, og andel af afdragsfrie realkreditlån er blevet udvalgt. Disse variable, som beskrevet i sektion 3.8, vil fremover blive anvendt og refereret til som følger: log_gdp, log_u, $afdragsfrihed$ samt husprisindekset som log_hp.

Inden en fejlkorrektionsmodel kan opstilles har det været nødvendigt at konkludere de anvendte variables integrationsorden for at sikre, at de er enten I(0) eller I(1), da dette er betingelserne for at kunne anvende en ARDL Bounds test. Yderligere er det også vigtigt at kende variablenes integrationsorden for at sikre, at der fremadrettet vil blive arbejdet med stationære serier, der ikke afhænger af tiden. Af tabel 10 fremgår resultaterne af de foretagede ADF-test for alle variable.

Variable	**Testværdi**	**Kritisk værdi**	**Integrationsorden**
log_hp	-1.791	-2.89	I(1) eller højere
log_gdp	-2.389	-3.45	I(1) eller højere
log_u	-3.448	-3.45	I(1) eller højere
$afdragsfrihed$	-1.623	-2.89	I(1) eller højere

Tabel 10: ADF-test for om variable er I(0) med tilhørende testværdier og kritiske værdier.

I alle tilfælde kan H_0-hypotesen om, at den undersøgte variabel skulle være integreret af

en højere orden end I(0) ikke afvises, og dermed er alle variable I(1) variable eller højere. Testen vil blive foretaget igen for at bekræfte, at variablene er integreret af første orden og ikke højere. Resultatet af denne test kan ses af tabel 11 herunder.

Variable	**Testværdi**	**Kritisk værdi**	**Integrationsorden**
log_hp	-3.095	-1.95	I(1)
log_gdp	-2.92	-2.89	I(1)
log_u	-3.38	-1.95	I(1)
$afdragsfrihed$	-1.964	-1.95	I(1)

Tabel 11: ADF-test for om variable er I(1) med tilhørende testværdier og kritiske værdier.

I alle tilfælde afvises h_0-hypotesen om, at den undersøgte variabel skulle være integreret af en højere orden end I(1). Dermed er alle variable I(1) variable, og er klar til anvendelse i fremtidige modeller.

Som det fremgår af sektion 3.4 tager ARDL Bounds testen udgangspunkt i en fejlkorrektionsmodel, der i dette projekt er opbygget således:

$$\begin{aligned}\Delta log_hp_t = &\sum_{i=1}^{n} \alpha_i \Delta log_hp_{t-i} + \sum_{i=1}^{n} \beta_i \Delta log_gdp_{t-i} + \sum_{i=1}^{n} \delta_i \Delta log_u_{t-i} \\ &+ \sum_{i=1}^{n} \gamma_i \Delta afdragsfrihed_{t-i} + \theta_1 log_hp_{t-1} + \phi_1 log_gdp_{t-1} \\ &+ \phi_2 log_u_{t-1} + \phi_3 afdragsfrihed_{t-1}\end{aligned} \tag{5.1}$$

Første del af modellen er den første difference af de uafhængige variable og fanger de kortsigtede effekter. Modellen blev testet ved at inkludere fire lags af hver uafhængig variabel. Dette valg blev truffet på baggrund af anvendelsen af kvartalsvise data samt ønsket om at fange så mange kortsigtede effekter som muligt. Samtidig blev antallet af lag begrænset til fire for at undgå at miste for mange frihedsgrader i modellen. Kun de signifikante lags er blevet beholdt i den endelige fejlkorrektionsmodel, resten er blevet fjernet ved hjælp af F-test. Alle niveauparametrene er ligeledes blevet fastholdt, da det er disse, der skal testes for kointegration med den afhængige variabel i ARDL Bounds testen. Da projektets hovedformål er at undersøge effekten af afdragsfrie realkreditlån på de reale huspriser kunne man dog alligevel være tilbøjelig til at fjerne de andre niveauparametre for kun at teste forholdet imellem disse to. Dog vil det udelade signifikante parametre fra

den endelige model, hvilket ændrer estimatet for afdragsfriheds variablen, skaber omitted variable bias og gør det fundne estimat ugyldigt.

Den endelig fejlkorrektionsmodel, der vil indgå i ARDL Bounds testen har følgende form:

$$\begin{aligned}\Delta log_hp_t = \alpha_1 \Delta log_hp_{t-1} + \alpha_4 \Delta log_hp_{t-4} + \delta_0 \Delta log_u_t + \delta_1 \Delta log_u_{t-1} \\ + \delta_2 \Delta log_u_{t-2} + \delta_4 \Delta log_u_{t-4} + \gamma_2 \Delta afdragsfrihed_{t-2} \\ + \theta_1 log_hp_{t-1} + \phi_1 log_gdp_{t-1} + \phi_2 log_u_{t-1} \\ + \phi_3 afdragsfrihed_{t-1}\end{aligned} \tag{5.2}$$

Denne model har bestået alle test for overholdelse af OLS antagelserne som vist tidligere i sektion 3.5. Estimationen af modellen kan ses af tabel 12

	Δlog_hp_t ARDL 1
Δlog_hp_{t-1}	0.614*** (0.091)
Δlog_hp_{t-4}	0.346*** (0.073)
Δlog_u_t	−0.034* (0.018)
Δlog_u_{t-1}	0.065*** (0.013)
Δlog_u_{t-2}	0.043*** (0.009)
Δlog_u_{t-4}	0.052*** (0.017)
$\Delta afdragsfrihed_{t-2}$	0.394*** (0.130)
log_hp_level	−0.062** (0.026)
log_gdp_level	0.023**

	(0.010)
log_u_level	−0.012
	(0.012)
afdragsfrihed_level	0.013*
	(0.008)
dummy_2008Q4	−0.028**
	(0.013)
N	95
Adjusted R^2	0.775

Noter: ***Signifikant ved et signifikansniveau på 1%.
**Signifikant ved et signifikansniveau på 5%.
*Signifikant ved et signifikansniveau på 10%.

Tabel 12: Estimeringen af fejlkorrektionsmodellen.

Nulhypotesen, at der ikke eksisterer et langsigtet forhold imellem variablene, ønskes testet imod alternativ hypotesen, at nogle af variablene som minimum har et langsigtet forhold:

$$H_0 \ : \ (\theta_1, \phi_1, \phi_2, \phi_3) = (0, 0, 0, 0)$$

$$H_1 \ : \ (\theta_1, \phi_1, \phi_2, \phi_3) \neq (0, 0, 0, 0)$$

Der vil blive foretaget en F-test for at teste signifikansen af niveauparametrene ved at sammenligne den fulde model med en begrænset model, hvor niveauparametrene vil blive sat til at være lig med 0. Den begrænsede model er derfor identisk med ligning 5.2 uden niveauparametrene og kan ses af ligning 5.3.

$$\begin{aligned} \Delta log_hp_t = \alpha_1 \Delta log_hp_{t-1} + \alpha_4 \Delta log_hp_{t-4} + \delta_0 \Delta log_u_t + \delta_1 \Delta log_u_{t-1} \\ + \delta_2 \Delta log_u_{t-2} + \delta_4 \Delta log_u_{t-4} + \gamma_2 \Delta afdragsfrihed_{t-2} \end{aligned} \tag{5.3}$$

Resultatet af F-testen fremgår af tabel 13, hvor der bliver opnået en F-statistik på 5.073.

Res.DF	**RSS**	**Df**	**Sum of Sq**	**F**	**Pr(>F)**
82	0.009	4	0.002	5.073	0.001

Tabel 13: F-test af fejlkorrektionsmodellen på den begrænsede model.

Fordi den anvendte model har tre kovariater og intet intercept skal den beregnede F-

statistik sammenlignes med de kritiske værdier fundet i Pesaran et al. (2001, s. 300) tabel CI (No intercept and no trend) linje 4 (K=3). De kritiske værdier er som følger:

- 5% signifikansniveau: Nedre grænse = 2.45; øvre grænse = 3.63. Fordi 5.073 > 3.34 kan nulhypotesen vedrørende et ikke eksisterende langsigtet forhold afvises.
- 1% signifikansniveau: nedre grænse = 3.42; øvre grænse = 4.84. Fordi 5.073 > 4.31 kan nulhypotesen vedrørende et ikke eksisterende langsigtet forhold afvises.

På både et 5% og 1% signifikansniveau kan nulhypotesen afvises, og der eksisterer derfor et langsigtet forhold imellem de undersøgte variable og deriblandt afdragsfrie realkreditlån og danske reale huspriser.

Dernæst vil den langsigtede effekt af variablen for afdragsfrihed blive beregnet ud fra følgende formel: $-\frac{\phi_3}{\theta_1}$ forklaret i sektion 3.4 for at finde ud af, hvilken effekt afdragsfrie realkreditlån har på lang sigt på de danske reale huspriser. Både ϕ_3 og θ_1 kan aflæses direkte i tabel 12 til hhv. 0.0134 og -0.062. Resultatet kan ses af ligning 5.4. Der er yderligere fortaget en Wald Chi-square test for at sikre, at det beregnede estimatet er forskelligt fra 0. Resultatet kan ses af tabel A.6 i appendiks og viser, at det beregnede estimat er signifikant på et 1% signifikansniveau.

$$-\frac{0.0134}{-0.062} = 0.217 \tag{5.4}$$

Ovenstående resultat skal fortolkes som en sammenhæng, hvor en stigning på én enhed i andelen af danske realkreditlån med afdragsfrihed medfører en stigning på 0.22% i de danske reale huspriser. Sammenfattende er det værd at bemærke, at den langsigtede effekt har et positivt fortegn, hvilket stemmer overens med tidligere litteraturs resultater om en positiv sammenhæng mellem udviklingen i afdragsfrie realkreditlån og de reale huspriser som beskrevet i sektion 2.

5.3 Robusthedstjek

Den undersøgte tidsperiode i ovenstående analyse strækker sig fra første kvartal 1994 til fjerde kvartal 2018, men da afdragsfrie realkreditlån først blev lovliggjort i Danmark fra fjerde kvartal 2003, så har denne dataserie som tidligere nævnt værdien 0 i alle perioder

op til førnævnte dato. For at sikre, at dette ikke har været problematisk for analysens resultater, er der yderligere blevet lavet en ARDL Bounds test på en fejlkorrektionsmodel med dataserier, der kun forløber sig fra fjerde kvartal 2003 til fjerde kvartal 2018. Dette er gjort med baggrund i at sikre robustheden af analysens resultater på trods af opbygningen af afdragsfriheds dataserien, hvor de mange observationer med værdien 0 kunne mindske eller forstørre den langsigtede effekt.

Den fulde estimering af modellen kan ses af tabel B.1 i appendiks. Det er værd at nævne, at denne model har bestået alle tests for overholdelse af OLS antagelserne, hvilket kan ses i appendiks tabel B.2-B.5. Fejlkorrektionsmodellen på det korte datasæt ser således ud:

$$\begin{aligned}\Delta log_hp_short_t = {} & \alpha_1 \Delta log_hp_short_{t-1} + \alpha_4 \Delta log_hp_short_{t-4} \\ & + \beta_3 \Delta log_gdp_short_{t-3} + \beta_4 \Delta log_gdp_short_{t-4} \\ & + \delta_0 \Delta log_u_short_t + \delta_1 \Delta log_u_short_{t-1} \\ & + \delta_2 \Delta log_u_short_{t-2} + \gamma_4 \Delta afdragsfrihed_short_{t-4} \\ & + \theta_1 log_hp_short_{t-1} + \phi_1 log_gdp_short_{t-1} \\ & + \phi_2 log_u_short_{t-1} + \phi_3 afdragsfrihed_short_{t-1}\end{aligned} \tag{5.5}$$

Der vil ligeledes blive foretaget en F-test for at sammenligne den korte model med dens tilhørende begrænsede model, hvor niveauparametrene vil blive sat til at være lig med 0. Den begrænsede model er derfor identisk med ligning 5.5 uden niveauparametrene og kan ses af ligning 5.6.

$$\begin{aligned}\Delta log_hp_short_t = {} & \alpha_1 \Delta log_hp_short_{t-1} + \alpha_4 \Delta log_hp_short_{t-4} \\ & + \beta_3 \Delta log_gdp_short_{t-3} + \beta_4 \Delta log_gdp_short_{t-4} \\ & + \delta_0 \Delta log_u_short_t + \delta_1 \Delta log_u_short_{t-1} \\ & + \delta_2 \Delta log_u_short_{t-2} + \gamma_4 \Delta afdragsfrihed_short_{t-4}\end{aligned} \tag{5.6}$$

Resultatet af F-testen fremgår af tabel 14, hvor der bliver opnået en F-statistik på 7.041.

Res.DF	**RSS**	**Df**	**Sum of Sq**	**F**	**Pr(>F)**
44	0.005	4	0.003	7.041	0.0002

Tabel 14: F-test af den korte fejlkorrektionsmodellen på den begrænset model.

Modellen testes på de samme grænseværdier fra Pesaran et al. (2001), da den har samme specifikationer:

- 5% signifikansniveau: Nedre grænse = 2.14; Øvre grænse = 3.34. Fordi 7.041 > 3.34 kan nulhypotesen vedrørende et ikke eksisterende langsigtet forhold afvises.
- 1% signifikansniveau: Nedre grænse = 2.82; Øvre grænse = 4.21. Fordi 7.041 > 4.31 kan nulhypotesen vedrørende et ikke eksisterende langsigtet forhold afvises.

På både et 5% og 1% signifikansniveau kan nulhypotesen afvises og der eksisterer derfor et langsigtet forhold imellem de undersøgte variable.

Igen vil den langsigtede effekt af afdragsfriheds variablen blive beregnet ud fra følgende formel: $-\frac{\phi_3}{\theta_1}$ for at finde frem til, hvilken langsigtet effekt afdragsfrie realkreditlån har på de danske reale huspriser ifølge modellen med den kortere tidshorisont. Både ϕ_3 og θ_1 kan aflæses direkte i tabel B.1 i appendiks til hhv. -0.146 og -0.073. Resultatet kan ses af ligning 5.7. Der er yderligere fortaget en Wald Chi-square test for at sikre, at det beregnede estimatet er forskelligt fra 0. Resultatet kan ses af tabel B.6 i appendiks og viser, at det beregnede estimat er signifikant på et 5% signifikansniveau.

$$-\frac{-0.146}{-0.073} = -1.991 \tag{5.7}$$

Der er en betydelig forskel imellem estimatet for den fulde periode og estimatet for den korte periode, hvilket kan være med til at vise, at de mange observationer med værdien 0 i afdragsfriheds dataserien ser ud til at have haft stor betydning for analysens konklusioner. Derfor validerer ovenstående tjek ikke de udarbejdede resultater og er dermed ikke med til at give rygdækning til, at afdragsfrie realkreditlån har en positiv effekt på de reale huspriser.

En overraskende observation fra det foretagne robusthedstjek er, at fortegnet på det langsigtede estimat er blevet ændret. Dette er ikke tidligere set i lignende studier, hvilket er med til at skabe tvivl om validiteten af ovenstående analyse. Dette kunne skyldes, at trend komponenten i afdragsfriheds variablen har stor betydning for udviklingen. Ved indførslen af afdragsfrihed blev der set en markant stigning de første år, hvilket kan skyldes, at variablen var ved at finde sit naturlige leje. Dette skete i en periode, hvor

huspriserne faldt markant under finanskrisen, hvorved de to tidsserier bevægede sig i hver sin retning. Spørgsmålet er, om huspriserne ville være faldet yderligere, hvis ikke der fra politisk side var blev indført muligheden for afdragsfrie realkreditlån et par år tidligere. Med andre ord kunne stigningen i afdragsfrihed som nævnt være trend specifik, og derfor vil data nu blive filtreret, så kun den cykliske komponent står tilbage for at undersøge udviklingen i tidsserien, der ikke kan tildeles trenden. Dette vil blive gjort ved hjælp af Hodrick-Prescott (HP) filteret, der er en af de mest kendte og mest brugte metoder til at separere trend fra den cykliske komponent af en dataserie (Hodrick og Prescott, 1997). Filteret er defineret som løsningen til optimeringsproblemet som fremgår af ligning 8.1 i appendiks, hvor y_t er den originale dataserie, τ_t er trend komponenten og c_t er den cykliske komponent. Metoden bygger på at minimere afvigelsen af den originale serie fra trenden, som er det første led, og krumningen af den estimerede trend, som er det andet led. Afvejningen imellem de to er givet ved udjævnings parameteren λ. Desto større λ er, desto mere jævn er den estimerede trend. (Hodrick og Prescott, 1997) (OECD, 2019)

Efter filtreringen er der kun den cykliske del af hver enkelt af de anvendte tidsserier tilbage, hvilket er formatet, der nu vil blive analyseret på.

5.4 OLS-estimering på filtreret data

HP-filterets udjævning af tidsserierne har betydet, at alle observationer med værdien 0 i afdragsfriheds variablen nu er forskellig fra 0 [2]. Derfor vil den forestående analyse anvende det korte datasæt fra fjerde kvartal 2003 til og med fjerde kvartal 2018 for at undgå potentielle problemer relateret hertil. Der vil til denne analyse blive anvendt de samme variable, som i de tidligere analyseafsnit.

Ligesom med de tidligere anvendte fejlkorrektionsmodeller er det igen vigtigt at fastslå integrations ordenen af de nye tidsserier, da en estimering ved hjælp af OLS kræver, at alle variable er stationære og dermed ikke har en middelværdi eller varians, der afhænger af tiden. Derfor er der blevet foretaget en stationaritetstest, hvis resultater kan aflæses af tabel 15.

[2] Dette kan give anledning til misvisende resultater.

Variable:	Testværdi	Kritisk værdi	Integrationsorden
log_hp	-2.402	-1.95	I(0)
log_gdp	-3.369	-1.95	I(0)
log_u	-2.462	-1.95	I(0)
$afdragsfrihed$	-3.096	-1.95	I(0)

Tabel 15: ADF-test for om de filtrerede variable er I(0) med tilhørende testværdier og kritiske værdier.

Resultaterne af stationaritets testen viser, at alle de anvendte variable er I(0) og dermed er stationære i niveau. Derfor kan der nu estimeres en OLS model for at undersøge de langsigtede effekter af den cykliske komponent af afdragsfrihed på den cykliske komponent af de reale huspriser. Modellen kan ses af ligning 5.8, hvor der er blevet anvendt de samme uafhængige variable, som i tidligere modeller plus fire lags af den afhængige variabel for at minimere seriekorrelation i modellen. Yderligere antages det også, at de reale huspriser i dag til dels afhænger af de reale huspriser i går, hvilket betyder, at der ligeledes er teoretisk grundlag for at inddrage dem i modellen. Valget af fire lags er igen sket på baggrund af det anvendte datagrundlag, der er kvartalsvise serier, hvorfor fire lags testes i modellen og alle beholdes, da disse er signifikante. Modellen har bestået alle tests af OLS antagelserne, hvis resultater kan ses i appendiks tabel C.1-C.4.

$$\begin{aligned} cycle_hp_t = cycle_hp_{t-1} + cycle_hp_{t-2} + cycle_hp_{t-3} + cycle_hp_{t-4} \\ + cycle_gdp_t + cycle_u_t + cycle_afdragsfrihed_t \end{aligned} \tag{5.8}$$

Ovenstående ligning har ført til estimering af nedenstående model:

	$cycle_log_hp_t$
$cycle_log_hp_{t-1}$	1.572^{***}
	(0.125)
$cycle_log_hp_{t-2}$	-1.125^{***}
	(0.212)
$cycle_log_hp_{t-3}$	0.784^{***}
	(0.214)
$cycle_log_hp_{t-4}$	-0.318^{**}
	(0.125)

$cycle_log_gdp_t$	-0.048
	(0.094)
$cycle_log_u_t$	0.027
	(0.017)
$cycle_afdragsfrihed_t$	0.410^{**}
	(0.154)
Observationer	57
R^2	0.946
Adjusted R^2	0.938
Standardfejl for fejlleddet	0.013 (df = 50)
F Statisstik	125.030^{***} (df = 7; 50)
Noter:	*p<0.1; **p<0.05; ***p<0.01

Tabel 16: Resultater af OLS estimering.

Vedrørende spørgsmålet angående effekten af afdragsfrie realkreditlåns betydning for udviklingen af de reale huspriser skal der kigges nærmere på det nederste estimat i ovenstående tabel. Estimatet på den cykliske del af afdragsfriheds variablen er 0.410 og er signifikant på et 5% signifikansniveau. Dette betyder, at en stigning på en enhed i den cykliske komponent af afdragsfriheds variablen vil medføre en stigning i den cykliske komponent af de reale huspriser på 0.4%. Ovenstående resultat betyder, at fjernes trend delen fra alle anvendte variable opnås en positiv effekt i stedet for en negativ effekt af afdragsfrie realkreditlån på de reale huspriser. Disse resultater stemmer i højere grad overens med tidligere studiers analyse af samme effekt. Samlet set præsenterer de forskellige modeller, der er blevet anvendt i denne sektion, blandede resultater. Disse resultater vil blive yderligere diskuteret i sektion 6, hvor validiteten af den foretagne analyse vil blive drøftet med fokus på fordele og ulemper ved den valgte tilgang.

6 Diskussion

Denne sektion vil forsøge at validere projektets resultater fra begge analysesektioner. Derfor vil der først gives en præsentation af de to DiD-modeller, og antagelserne i et DiD-framework vil blive diskuteret med henblik på at vurdere, hvorvidt disse er overholdt. Derefter vil de fundne estimater i den økonometriske analyse blive diskuteret med henblik på, om disse var forventelige ift. den undersøgte relation imellem de reale huspriser og implementeringen af afdragsfrie realkreditlån i Danmark. Afslutningsvis vil de opnåede resultater blive inddraget i en politisk kontekst for at undersøge de samfundsøkonomiske konsekvenser heraf.

6.1 Validering af resultater

6.1.1 DiD-analysen

I DiD-frameworket er SCM blevet anvendt til at opbygge en passende kontrolgruppe til Danmark. Når der arbejdes med DiD-analyser, er den centrale antagelse, at der er parallelle trends mellem de to grupper i perioden op til behandlingen. Dette har kun i nogen grad været tilfældet (se figur 8).

Den manglende overholdelse af antagelsen kan skyldes flere begrænsninger, som ligger i processen. Til udvælgelsen af donor-lande til den syntetiske kontrolgruppe er der arbejdet ud fra et tilgængelighedsprincip, da det er meget sparsomt med lande, der ikke har haft afdragsfrie lån i nogen grad. Derfor er lande med lav anvendelse af denne også godtaget til analysen. Der er yderligere fokuseret på europæiske lande, som teoretisk burde ligne Danmark mest, men de nordiske lande har enten haft eller har stadig afdragsfrie lån. Dette afspejler sig også i, at de højest vægtede lande i den syntetiske kontrolgruppe er Italien og Spanien med en vægt på hhv. 42.2% og 38.5%. Disse lande er objektivt udvalgt i analysen som lande der ligner Danmark ud fra anvendte makroøkonomiske variable (udviklingen i BNP-indekset har hertil været vægtet højest fra modellens side). Det kan argumenteres, at disse lande ikke ligner Danmark i høj grad nok til at danne en parallel trend i huspriser, men det er dog de tilgængelige lande, der har haft mindst (eller ingen) grad af afdragsfrie lån, se figur 7.

Endnu en usikkerhed til kontrolgruppen er, at det til estimering af makrovariablene

tilhørende kontrolgruppen ligeledes har været nødvendigt at udregne disse med de vægte opnået gennem Synth-pakken, som er anvendt til at lave den syntetiske kontrolgruppe. Dette øger usikkerheden vedrørende kontrolgruppen, da det i stedet for blot at være outcome variablen, der er en estimeret tidsserie, nu også er tre andre makroøkonomiske variable, der er estimerede. Således er det samlet fire estimerede tidsserier, der anvendes i regressionerne.

Denne pakke er ydermere ikke designet til at arbejde med kvartalsvise datasæt. Dette kan have haft en effekt, da treatment skete i fjerde kvartal 2003 - i stedet behandler analysen hele år 2003 som treatment året. Dog kan det argumenteres, at modellerne betragter årene efter 2003 (2004-2018) som post-treatment perioden, hvorfor anvendelsen af kvartalsvise eller årlige data ikke burde have stor indflydelse på resultaterne udover muligvis at udjævne nogle effekter.

I første SDiD-model blev regressionen $hp = gruppe + tid + did$ opstillet. Denne gav et positivt estimat for DiD-estimatoren, hvilket er i tråd med litteraturen og projektets økonometriske analyse. Modellen var dog kun signifikant ved et signifikansniveau på 10%, hvilket taler imod modellens validitet. Usikkerheden i modellens estimater kan skyldes confounding factors, altså en form for omitted variable bias, hvor der har været en forskel i udviklingen af en variabel (der har effekt på huspriser) i treatmentgruppen ift. kontrolgruppen i post-treatment perioden.

For at tage højde for nogle af disse confounding factors er der blevet opstillet en DiD-model med kontrolvariable. Heri blev mulige confounding factors som BNP, arbejdsløsheden og rentesatsen implementeret direkte i modellen. Dette forbedrede i høj grad signifikansniveauet for DiD-estimatoren, og de implementerede confounding variable viste ligeledes signifikante estimater. Estimatet for DiD kan virke højt ift. de andre implementerede variable, men her skal det understreges, at DiD-estimatoren er en gennemsnitlig samlet effekt af implementeringen af afdragsfrie realkreditlån siden 2003, mens estimatet for fx BNP er en enheds stigning i BNP-indekset mod en stigning i husprisindekset. DiD-estimatet påstår, at de afdragsfrie realkreditlån har hævet husprisindekset med 18 indekspoint i perioden fra 2003-2018. I denne periode er husprisindekset steget fra 78.6 til 115.3 (i reale priser). Derved skyldes 49% af denne stigning, ifølge kontrolvariable-

modellen, implementeringen af afdragsfrie realkreditlån. Dette tal forudsætter dog, at ingen af de anvendte lande i kontrolgruppen har oplevet nogen form for afdragsfrie realkreditlån, hvilket som nævnt kan være en naiv tankegang.

Teoretisk var estimaterne for alle de implementerede variable udover renten forventelige. Empirisk set har BNP og huspriserne haft en positiv korrelation, da et land i økonomisk vækst øger indkomsten og dermed købekraften og efterspørgslen. Således er BNP-estimatet også positivt. Arbejdsløsheden og renten må forventes at have en negativ effekt på huspriserne. Arbejdsløsheden er negativt korreleret med BNP og giver lavere indkomstniveauer, hvilket må sænke huspriserne. Hvis renten stiger, vil det blive dyrere at finansiere sit boliglån, hvilket alt andet lige må sænke efterspørgslen efter boliger og derigennem huspriserne. Derfor kan estimatet for renten ses som værende misvisende, da dette var positivt. Således kan renten i modellen måske have fanget en forkert effekt og derigennem opfanget noget af DiD-estimatorens effekt. Dog var renten også den eneste variabel i robusthedstjekket, der ikke var signifikant.

For at validere denne kontrolvariable-model blev der som nævnt udført en robusthedsanalyse, som ikke havde den store effekt på signifikansen af estimaterne. DiD-esimatoren var stadig signifikant, omend estimatet hertil blev lidt lavere. Dette vurderes ikke at have den store effekt for den endelige konklusion af modellerne, da post-treatment i den kortere tidsperiode kun strækker sig over 8 år (hvoraf finanskrisen også har haft en effekt på flere af disse år). Derudover så Danmark en større og hurtigere stigning i husprisindekset fra 2010-2018 end kontrolgruppen gjorde, hvorfor der helt naturligt vil være større estimater for modellerne med de længere tidsserier. Samlet set understøttes validiteten af kontrolvariable-modellen med det forbehold, at den syntetiske kontrolgruppe ikke lever helt op til antagelsen om parallelle trends i et DiD-framework.

6.2 Fejlkorrektionsmodeller

I projektets økonometriske analyse findes signifikante resultater for en langsigtet sammenhæng imellem de reale huspriser og andelen af husholdninger, der anvender afdragsfrie realkreditlån på dansk data. Det kan ses i den anvendte CUSUM-test, som er vist i appendiks figur A.3, at recursive CUSUM-testen kan give problematiske resultater. Dog blev det tidligere konkluderet i projektet, at dette ikke var tilfældet, hvorfor der kan siges ikke

at være sket et signifikant skifte i estimaterne i den anvendte fejlkorrektionsmodel, der blev præsenteret i sektion 5. Af denne årsag vurderes det, at der findes tidskonsistente estimater på baggrund af dansk data for perioden første kvartal 1994 til og med fjerde kvartal 2018. Dette medgiver, at anvendelsen af en dataserie for andelen af husholdninger, der anvender afdragsfrie realkreditlån, hvor der indgår 39 observationer med værdien 0 i starten af serien, ikke har haft en indvirkning på tidskonsistensen af analysens resultater. Det beregnede positive estimat fra den økonometriske analyse stemmer yderligere overens med tidligere studiers lignende fund af en positiv sammenhæng mellem de undersøgte variable jf. Karpestam og Johansson (2019) og Barlevy og Fisher (2021).

For yderligere at validere analysens estimat blev endnu en økonometrisk analyse foretaget med en kortere tidshorisont. Denne kortere tidshorisont skulle sikre, at dataserien for andelen af husholdninger, der anvender afdragsfrie realkreditlån, hvori der indgik 39 observationer med værdien 0, ikke havde betydning på størrelsen af estimatet og dermed effekten af sammenhængen mellem den nævnte variabel og de reale huspriser. Resultatet af denne analyse viste dog, at ved kun at ændre længden på de anvendte tidsserier, skiftede den undersøgte sammenhæng fortegn, hvorved en stigning i brugen af afdragsfrie realkreditlån betød et fald i de reale huspriser. Denne ændring var højst uventet omend stadig interessant, da det kunne betyde, at de mange observationer med værdien 0 i afdragsfriheds variablen alligevel havde en betydning ikke kun for størrelsen, men også for retningen af det undersøgte forhold. Ændringen i fortegnet på det undersøgte estimat kan også skyldes andre forhold end ovenstående, såsom generelt kortere tidsserier med færre observationer, der gør analysen mindre robust ved at undersøge en kortere tidsperiode, hvor der kun tages højde for effekter, der er sket i den anvendte periode. Dog optræder variable, hvor serien starter med x antal observationer med værdien 0 ofte i økonometriske analyser i form af dummy-variable. Deres tilstedeværelse giver ikke grobund for biased estimater eller lignende, hvilket taler for, at det heller ikke skulle være et problem for det beregnede estimat i denne analyse trods de mange nul observationer.

Det er også relevant at bemærke, at selvom afdragsfrie realkreditlån har eksisteret i 20 år og derfor er et veletableret finansielt instrument, så er det stadig et relativt ungt koncept. Dette betyder, at forskellige trends, udsving og sammenhænge endnu ikke er fuldt ud etablerede, og det kan være svært at undersøge dem fuldt ud på nuværende tidspunkt.

Med tiden kan det dog blive lettere at undersøge disse forhold, når der dannes yderligere empiri og en længere tidsserie af data er tilgængelig. Her tænkes der fx på, at starten af analyseperioden med den korte tidshorisont i fjerde kvartal 2003 angiver første gang, at det er muligt for de danske husholdninger at anskaffe sig realkreditlån med afdragsfrihed. Dette medfører, at tidsserien vil bruge de første år, måske endnu længere, på at finde sit naturlige niveau, hvorved indførselsperioden i 2003 kan være problematisk at inddrage i en analyse grundet den markante stigning op imod et naturligt niveau, hvor serien efterfølgende vil forløbe sig ud fra. Imens denne stigning i afdragsfrihed fandt sted faldt den afhængige variabel, de reale huspriser, markant under finanskrisen, hvilket kan give anledning til at tro, at de to serier er negativt korreleret. Dette kan teoretisk være svært at finde beviser for. Ovenstående observation gav anledning til at udføre en økonometrisk undersøgelse, hvor trenden i de undersøgte variable ikke indgik for udelukkende at se på de cykliske udsving i serierne. Denne model viste igen en positiv sammenhæng imellem andelen af husholdninger, der anvender afdragsfrie realkreditlån og de reale huspriser. Det er vigtigt at nævne, at dette er en positiv sammenhæng imellem de cykliske komponenter af de to serier, hvorfor resultaterne ikke kan sammenlignes en til en med de tidligere økonometriske modeller. Dog kan retningen af forholdet anvendes til yderligere at understøtte resultaterne fra den første analyse med den lange tidsserie, der fandt en lignende sammenhæng.

6.3 Politisk perspektiv og relevans

Den foregående diskussion af projektets resultater giver mulighed for at integrere disse i en politisk og samfundsøkonomisk kontekst, hvilket er afgørende for at vurdere anvendeligheden af de udførte analyser samt forstå, hvordan politiske interventioner kan påvirke boligmarkedet.

Forskellige politiske beslutninger og tiltag kan påvirke udbud og efterspørgsel samt prisniveauerne på boliger. Fx kan ændringer i skattelovgivningen, regulering af lånemuligheder eller ændringer i boligstøtteprogrammer påvirke incitamentet til at købe eller sælge boliger ved at øge husholdningernes købekraft eller lempe investeringsmulighederne. Desuden kan politiske beslutninger omkring rentesatser og centralbankpolitik have en afgørende indflydelse på boligmarkedet. Ændringer i renteniveauerne kan påvirke bolig-

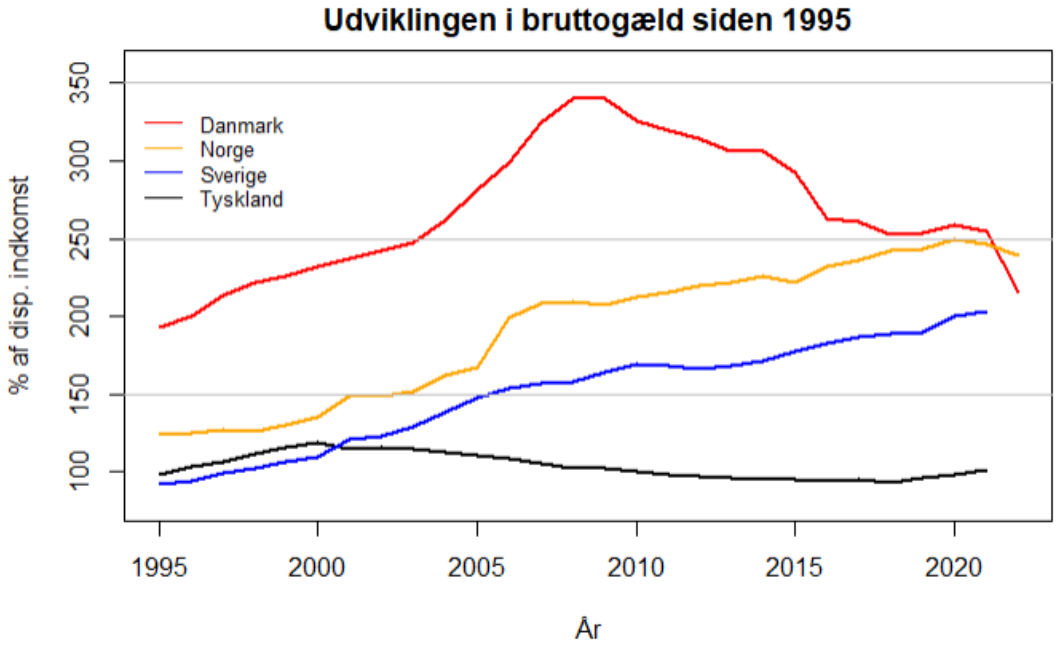
prisernes bevægelser samt adgangen til finansiering af boligkøb. Samlet set kan politisk indblanding have både kortsigtede og langsigtede konsekvenser for boligmarkedet. Det er vigtigt for politiske beslutningstagere at afveje og vurdere disse konsekvenser nøje for at sikre en stabil og bæredygtig udvikling på boligmarkedet.

I 2003 blev der indført politiske tiltag på boligmarkedet, der tillod afdragsfrie realkreditlån. Disse tiltag gav de danske husholdninger mere fleksibilitet i deres realkreditlån, hvor afdragsfriheden kunne bruges i perioder med tabt indkomst eller for unge med en mindre indkomst til at komme ind på boligmarkedet. Omvendt bragte den også negative effekter med sig i form af øget sandsynlighed for, at realkreditgælden ikke blev nedskrevet, hvilket stiller husholdningerne i udsatte situationer, når boligpriserne falder. Projektets resultater understøtter eksistensen af en positiv sammenhæng mellem boligpriserne og anvendelsen af afdragsfrie realkreditlån, hvilket kan betyde, at der uden implementeringen fra politisk side ville være sket et endnu større fald i boligpriserne end hvad der egentligt var realiteten under finanskrisen. På den anden side, kan det også argumenteres for, at boligboblen ikke var vokset i så høj en grad, hvis de afdragsfrie realkreditlån ikke var blevet introduceret. Dette er med til at fastslå sådanne politiske indvendingers betydning for boligmarkedet som helhed.

Indførelsen af afdragsfrihed i 2003 gjorde yderligere den omkostningseffektive belåning af ejerboliger i det danske realkreditsystem stærkere. Dette medførte, at danske husholdningers bruttogæld (gælden uden hensyntagen til eventuel formue) blev markant forhøjet sammenlignet med andre OECD medlemslande, jf. figur 10. Allerede før indførslen af afdragsfrihed var Danmark det land, som havde den største bruttogæld set ift. den disponible indkomst med 192.80% i 1995. Siden da er husholdningernes gældskvote steget fra 192.80% af den disponible indkomst til 339.78% i 2008 og siden faldet til 215.12% i 2022. Hele udviklingen kan ses af figur 10.(Finansministeriet, 2014)

Den øgede bruttogæld siden 1995 kan delvist forklares af faldet i realkreditrenterne, hvis udvikling kan ses af tabel E.1 i appendiks. Lavere renter gør det muligt for husholdningerne at forrente og afdrage en større gæld baseret på deres indkomstniveau. Samtidig betyder det lavere renteniveau, at det har været mere fordelagtigt for husholdningerne at flytte deres forbrug fremad i tiden, da gevinsten ved opsparing er reduceret. Omvendt kan det nuværende fald i bruttogælden tilskrives de stigende realkreditrenter, der alt andet

lige formindsker størrelsen af gæld, der kan optages på samme indkomstniveau.



Figur 10: Udviklingen i bruttogælden for udvalgte lande.
Kilde: (OECD.data, 2022a).

I perioden 2001-2004, inden de store stigninger i boligpriserne i 2005-2006, lå den gennemsnitlige belåningsgrad (dvs. restgælden ift. boligens værdi) på ca. 45%, men steg de efterfølgende år til godt 60%. Stigningen i belåningsgraden er sket på baggrund af forskellige faktorer, herunder det store prisfald på ejendomme i 2009 efter finanskrisen og det yderligere fald i 2012 i forbindelse med gældskrisen i visse EU-lande. Derudover kan indførelsen af afdragsfrie lån siden 2003 have bidraget til en strukturel stigning i belåningsgraden, da husholdningerne dermed har fået lettere adgang til at reducere deres opsparing i boligen uden at skulle optage ekstra lån.(Finansministeriet, 2014)

Ved den seneste opgørelse foretaget af Danmarks statistik (2022) er belåningsgraden igen nede omkring de 45%, hvilket betyder, at de danske husholdninger er bedre stillet under en eventuel krise, end de var i 2008 ved en højere belåningsgrad.

7 Konklusion

Projektet har forsøgt at besvare hvilken betydning implementeringen af afdragsfrie realkreditlån i Danmark har haft på de reale huspriser. Hertil har anvendelsen af den syntetiske Difference-in-Differences metode muliggjort en analyse af effekten af implementeringen af afdragsfrie realkreditlån i Danmark i 2003 på boligpriserne. Ved at konstruere en syntetisk kontrolgruppe baseret på paneldata for 6 donor-lande og tildele vægte baseret på ligheden mellem Danmark og de udvalgte lande blev det muligt at vurdere virkningerne af afdragsfrie realkreditlån på boligprisindekset. Resultaterne af analysen indikerer en positiv effekt af implementeringen af afdragsfrie realkreditlån på husprisindekset i Danmark. Den beregnede syntetiske kontrolgruppe viser, at boligprisindekset i Danmark steg hurtigere efter 2003 ift. den syntetiske kontrolgruppe. Både den oprindelige SDiD-analyse og en model med kontrolvariable bekræfter en statistisk signifikant positiv sammenhæng mellem implementeringen af afdragsfrie lån og stigningen i husprisindekset.

Relationen blev ydermere undersøgt ved hjælp af en ARDL Bounds test, hvor det blev påvist, at afdragsfrie realkreditlån har haft en signifikant effekt på udviklingen af de danske reale huspriser. Den beregnede effekt viste, at en stigning på 1% i andelen af afdragsfrie realkreditlån fører til en stigning på 0.23% i de danske reale huspriser på lang sigt. Dette resultat bekræfter tidligere litteratur, der har fundet en positiv sammenhæng mellem afdragsfrie realkreditlån og huspriser. Samme analyse blev ligeledes foretaget på et datasæt med en kortere tidshorisont, hvis resultat viste et negativt estimat for samme sammenhæng. Dette blev formodet at kunne skyldes trend komponenten i de kortere tidsserier, hvorfor disse blev fjernet, og kun de cykliske udsving i den undersøgte relation blev anvendt. Her fandtes ligeledes et positivt estimat for den undersøgte sammenhæng, hvor en stigning på en enhed i den cykliske komponent af afdragsfriheds variablen vil medføre en stigning i den cykliske komponent af de reale huspriser på 0.4%.

Projektet har ligeledes prøvet at inddrage det undersøgte forhold i en politisk sammenhæng, da forskellige politiske tiltag og beslutninger har betydelig indflydelse på boligmarkedet, herunder ændringer i skattelovgivningen, regulering af lånemuligheder, boligstøtteprogrammer og rentesatser. Disse politiske indgreb har været med til at påvirke udbud, efterspørgsel og prisniveauer på boliger både på kort og lang sigt. Det er vigtigt

for politiske beslutningstagere at afveje og vurdere konsekvenserne nøje for at sikre en stabil og bæredygtig udvikling på boligmarkedet.

Indførelsen af afdragsfrie realkreditlån i 2003 i Danmark har haft både positive og negative virkninger. Projektets resultater understøtter eksistensen af en positiv sammenhæng mellem brugen af afdragsfrie lån og boligprisernes udvikling. Det tyder på, at boligpriserne ville have oplevet et endnu større fald under finanskrisen, hvis de afdragsfrie lån ikke var blevet indført. Samtidig kan det også argumenteres for, at boligboblen ikke ville have nået samme omfang uden disse lån. Dette understreger betydningen af politiske indgreb på boligmarkedet.

Den danske husholdningsgæld er steget markant i perioden 1995-2009, delvist som følge af faldende realkreditrenter, der har gjort det muligt at forrente og afdrage en større gæld baseret på indkomstniveauet. I de senere år er bruttogælden faldet på grund af stigende realkreditrenter, der begrænser den mængde gæld, der kan optages på samme indkomstniveau.

Dette projekt har undersøgt hvilken betydning implementeringen af afdragsfrie realkreditlån i Danmark har haft på de reale huspriser. Samlet set kan det konkluderes, at projektet finder en positiv sammenhæng mellem denne implementering og de reale huspriser i Danmark. Dette bidrager med indsigt i betydningen af politiske tiltag på boligmarkedet og understreger behovet for yderligere forskning og analyser for at få en dybere forståelse af de komplekse faktorer, der påvirker denne sammenhæng fuldt ud.

8 Litteratur

Abadie et al., **2011**. Alberto Abadie, Alexis Diamond og Jens Hainmueller. *Synth: An R package for synthetic control methods in comparative case studies*. Journal of Statistical Software, 42(13), 1–17, 2011. ISSN 15487660. doi: 10.18637/jss.v042.i13.

Abdi og Williams, **2010**. Hervé Abdi og Lynne J. Williams. *Principal component analysis*. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433–459, 2010. ISSN 19395108. doi: 10.1002/wics.101. URL `https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wics.101`.

Kim Abildgren. A Chart & Data Book on the Monetary and Financial History of Denmark. 2017.

Alves, **2022**. Matheus Facure Alves. *25 - Synthetic Difference-in-Differences*, 2022. URL `https://matheusfacure.github.io/python-causality-handbook/25-Synthetic-Diff-in-Diff.html`.

Dmitry Arkhangelsky, Susan Athey, Stefan Wager, David A. Hirshberg og Guido W. Imbens. SYNTHETIC DIFFERENCE IN DIFFERENCES. URL `https://www.proquest.com/docview/2266420309?accountid=8144`. 2016.

Bäckman og Lutz, **2020**. Claes Bäckman og Chandler Lutz. *The impact of interest-only loans on affordability*. Regional science and urban economics, 80, 103376, 2020. ISSN 0166-0462. doi: 10.1016/j.regsciurbeco.2018.06.004.

Barlevy og Fisher, **2021**. Gadi Barlevy og Jonas D.M. Fisher. *Why were interest-only mortgages so popular during the U.S. housing boom?* Review of economic dynamics, 41, 205–224, 2021. ISSN 1094-2025. doi: 10.1016/j.red.2020.09.001.

BIS, **2023**. BIS. *Real residential property prices*, 2023. URL `https://stats.bis.org/statx/srs/table/h2?m=628&p=20224&c=`.

Brownlee, **2020**. Jason Brownlee. *A Gentle Introduction to k-fold Cross-Validation*, 2020. URL `https://machinelearningmastery.com/k-fold-cross-validation/`.

Columbia University, **2013**. Columbia University. *Difference-in-Difference Estimation*, 2013. URL `https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation`.

Dam et al., **2011**. Niels Arne Dam, Tina Saaby Hvolbøl, Erik Haller Pedersen, Peter Birch Sørensen og Susanne Hougaard Thamsborg. *Boligboblen der bristede: Kan boligpriserne forklares? Og kan deres udsving dæmpes?* Danmarks Nationalbank Kvartalsoversigt 1. kvartal Del 1, pages 47–70, 2011. URL `https://www.nationalbanken.dk/da/publikationer/Sider/2011/03/Kvartalsoversigt-2011,-1.-kvartal,-del-1.aspx`.

Danmarks statistik, **2022**. Danmarks statistik. *EJERFOF1 - Husholdningernes formue i fast ejendom og gældskomponenter for familier*, 2022.

Danmarks Statistik, **4 2023a**. Danmarks Statistik. *NKN1: Forsyningsbalance, Bruttonationalprodukt (BNP), beskæftigelse mv. efter transaktion, prisenhed og sæsonkorrigering*, 2023a.

Danmarks Statistik, **2023b**. Danmarks Statistik. *DNPIN: Indlån i pengeinstitutter i alt efter instrument, datatype, sektor i ind- og udland, land og valuta*, 2023b.

Danmarks Statistik, **2023c**. Danmarks Statistik. *FORV1: Forbrugerforventninger (nettotal) efter indikator*, 2023c.

Department of Housing og Government, **2017**. Planning Department of Housing og Local Government. *Mortgage market statistics*, 2017. URL `https://www.gov.ie/en/collection/1c230-mortgage-market-statistics/`.

Dickey og Fuller, **1979**. David A. Dickey og Wayne A. Fuller. *Distribution of the Estimators for Autoregressive Time Series With a Unit Root*. Journal of the American Statistical Association, 74(366), 427, 1979. ISSN 0162-1459. doi: 10.2307/2286348.

Enders, **2014**. Walter. Enders. *Applied econometric time series*. Wiley, Hoboken, NJ, fourth edition. edition, 2014. ISBN 9781118808566.

Eriksen, **2020**. Jesper Eriksen. *Long-run relationships and ARDL Bounds testing*, Aalborg Universitet, 2020. URL https://www.moodle.aau.dk/pluginfile.php/2140019/mod_resource/content/1/Example_ARDL.pdf.

European Mortgage Federation, **5 2019**. European Mortgage Federation. *Interest-only loans study*, European Mortgage Federation, 5 2019.

Finans Danmark, **2023**. Finans Danmark. *Obligationsrenter*, 2023. URL https://finansdanmark.dk/tal-og-data/boligstatistik/obligationsrenter/.

Finansinspektionen, **2017**. Finansinspektionen. *The Swedish Mortgage Market (2017)*, Finansinspektionen, 2017.

Finansministeriet, **2014**. Finansministeriet. *Privat opsparing og gæld*, finansministeriet, 2014. URL https://fm.dk/media/13883/kapitel2_finansredegoerelse2014_web.pdf.

Finlands Bank, **2023**. Finlands Bank. *Loans and deposit - Housing loans*, 2023. URL https://www.suomenpankki.fi/en/statistics2/dashboards-and-open-data/loans-and-deposit/#loans_and_deposits.

Frank van Hoenselaar et al., **2021**. Frank van Hoenselaar, Boris Cournède, Federica De Pace og Volker Ziemann. *Mortgage finance across OECD countries*. IDEAS Working Paper Series from RePEc, (1693), 1–36, 2021. doi: 10.1787/f97d7fe0-en.

Glen, **a**. Stephanie Glen. *Lasso Regression: Simple Definition*, a. URL https://www.statisticshowto.com/lasso-regression/.

Glen, **b**. Stephanie Glen. *Regularization: Simple Definition, L1 & L2 Penalties*, b. URL https://www.statisticshowto.com/regularization/.

Gregory, **2022**. Jesse Gregory. *Economics 695 : Difference in Differences: Introduction*, 2022.

Hodrick og Prescott, **1997**. Robert J. Hodrick og Edward C. Prescott. *Postwar U.S. Business Cycles: An Empirical Investigation*. Journal of money, credit and banking, 29(1), 1–16, 1997. ISSN 0022-2879. doi: 10.2307/2953682.

Karpestam og Johansson, **2019**. Peter Karpestam og Sebastian Johansson. *Interest-only-mortgages and housing market fluctuations in Denmark*. Journal of housing economics, 46, 101627, 2019. ISSN 1051-1377. doi: 10.1016/j.jhe.2019.03.004.

Keita, **2023**. Zoumana Keita. *Principal Component Analysis in R Tutorial*, 2023. URL `https://www.datacamp.com/tutorial/pca-analysis-r`.

Lionsgate Capital, **2023**. Lionsgate Capital. *Types of mortgages in Spain*, 2023. URL `https://lionsgatecapital.com/types-of-mortgages-in-spain`.

Nationalbankens Statistikbank, **2023**. Nationalbankens Statistikbank. *DNRUDDKI: Indenlandsk realkreditudlån fra realkreditinstitutter efter instrument, datatype, indenlandsk sektor, valuta, ejendomskategori, variabelt eller fast forrentet og afdrag*, 2023.

Nationalbankens Statistikbank, **2013a**. Nationalbankens Statistikbank. *DNEJER: Realkredit udlån til indlændinge fordelt på ejendomskategori, lånetype og løbetid*, 2013a.

Nationalbankens Statistikbank, **2013b**. Nationalbankens Statistikbank. *DNINDL1: Pengeinstitutternes indlån og udstedte gældsinstrumenter fordelt på løbetid*, 2013b.

Nguyen, **2023**. Mike Nguyen. *A Guide on Data Analysis*. page 26 Synthetic Control, 2023. URL `https://bookdown.org/mike/data_analysis/synthetic-control.html`.

OECD, **2019**. OECD. *Trend estimation method*, 2019. URL `https://www.oecd-ilibrary.org/sites/361fabc3-en/index.html?itemId=/content/component/361fabc3-en`.

OECD.data, **2022a**. OECD.data. *Household debt*, 2022a.

OECD.data, **2022b**. OECD.data. *Housing prices*, 2022b.

Økonomi - og erhvervsministeriet, **9 2008**. Økonomi - og erhvervsministeriet. *Afdragsfrie lån og prisstigninger på boligmarkedet*, Økonomi - og erhvervsministeriet,

9 2008. URL https://www.ft.dk/samling/20081/almdel/p%C3%B8u/spm/2/svar/575074/602102.pdf.

Oxford Reference. Oxford Reference. *Scree test*. URL https://www.oxfordreference.com/display/10.1093/oi/authority.20110803100449263;jsessionid=ADFB78356209C5A260972BC891F992AF.

Pesaran et al., **2001**. M. Hashem Pesaran, Yongcheol Shin og Richard J. Smith. *Bounds testing approaches to the analysis of level relationships*. Journal of applied econometrics (Chichester, England), 16(3), 289–326, 2001. ISSN 0883-7252. doi: 10.1002/jae.616.

Pfaff, **2008**. Bernhard Pfaff. *Analysis of Integrated and Cointegrated Time Series with R*. Use R! Springer New York, New York, NY, 2 edition, 2008. ISBN 9780387759661. doi: 10.1007/978-0-387-75967-8.

Rasmussen, **3 2021**. Jørgen Munksgaard Rasmussen. *Efter store prisstigninger: Finansiel vagthund vil dæmpe boligpriser*, 2021. URL https://www.bolius.dk/efter-store-prisstigninger-finansiel-vagthund-vil-daempe-boligpriser-96501.

Realkredit Rådet, **2015**. Realkredit Rådet. *Den klassiske danske realkreditmodel*, 2015. URL https://finansdanmark.dk/media/yurpkl45/den-danske-realkreditmodel.pdf.

Realkreditrådet, **2007**. Realkreditrådet. *Låntagernes brug af konverteringsgevinster og provenu fra tillægslån*, ealkreditrådet, 2007.

Statistics Norway, **2023**. Statistics Norway. *Banker og kredittforetak*, 2023. URL https://www.ssb.no/statbank/table/09562/.

Totalkredit, **2023**. Totalkredit. *Værd at vide om afdragsfrie lån*, 2023. URL https://www.totalkredit.dk/boliglan/kob-af-bolig/bliv-klogere-pa-realkredit/laan-med-eller-uden-afdrag/#er-lan-uden-afdrag-mere-kursfolsomme-end-lan-med-afdrag.

Wooldridge, **2018**. Jeffrey M. Wooldridge. *Introductory Econometric*. Cengage Learning, 6th edition, 2018. ISBN 978-1-305-27010-7.

Yiu, **2019**. Tony Yiu. *The Curse of Dimensionality*, 2019. URL `https://towardsdatascience.com/the-curse-of-dimensionality-50dc6e49aa1e`.

Appendiks

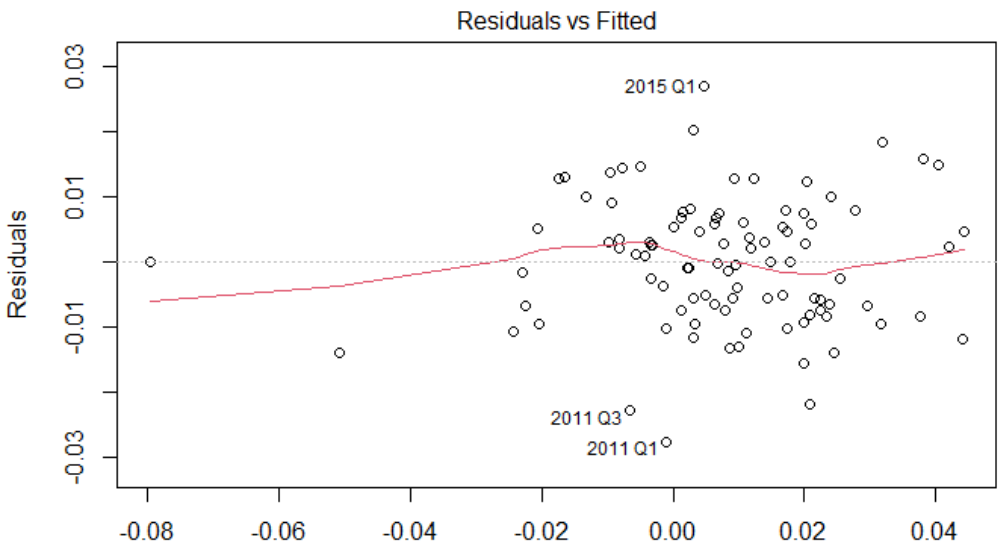
Appendiks A: Den overordnede fejlkorrektionsmodel

RESET	df1	df2	p-value
2.334	2	81	0.1034

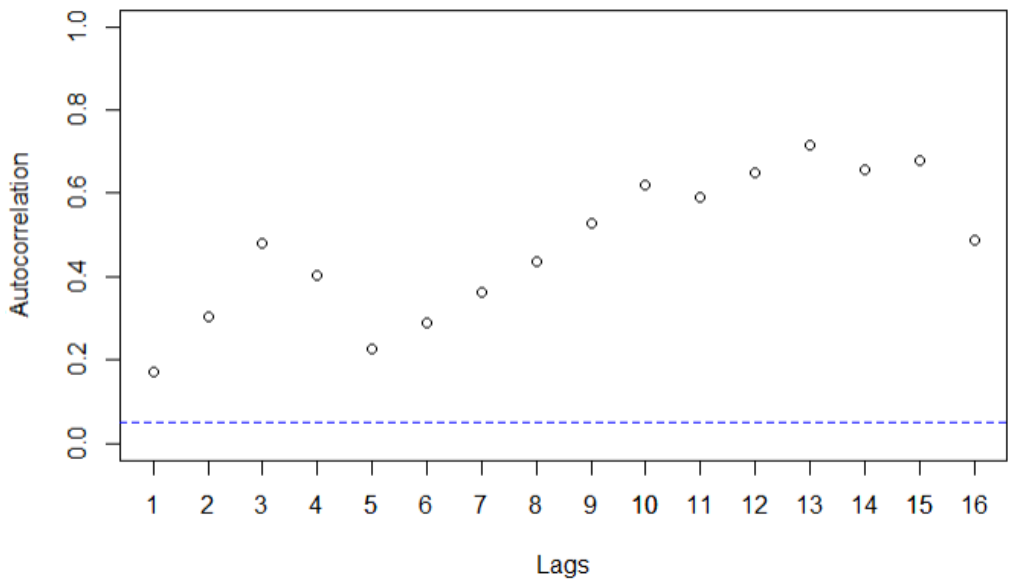
Tabel A.1: RESET-test for test af lineære parametre.

Variable	gdp	c	u	mb	gb	s	b	andel	tillid
gdp	-								
c	0.97	-							
u	-0.85	-0.797	-						
mb	-0.894	-0.877	0.69	-					
gb	-0.894	-0.876	0.723	0.985	-				
s	0.888	0.918	-0.698	-0.893	-0.914	-			
b	-0.374	-0.433	0.365	0.464	0.523	-0.637	-		
andel	0.798	0.86	-0.662	-0.796	-0.82	0.948	-0.657	-	
tillid	-0.049	-0.057	0.319	-0.035	0.058	-0.058	0.271	-0.114	-

Tabel A.2: Korrelationstest af alle medvirkende variable.



Figur A.1: Fitted værdier vs. fejlleddet for den anvendte fejlkorrektionsmodel.



Figur A.2: ARCH-test for autokorrelation på 16 lags - Den kombinerede p-værdi over box-testen er 0.58.

BP	Df	P-værdi
6.0644	12	0.9144

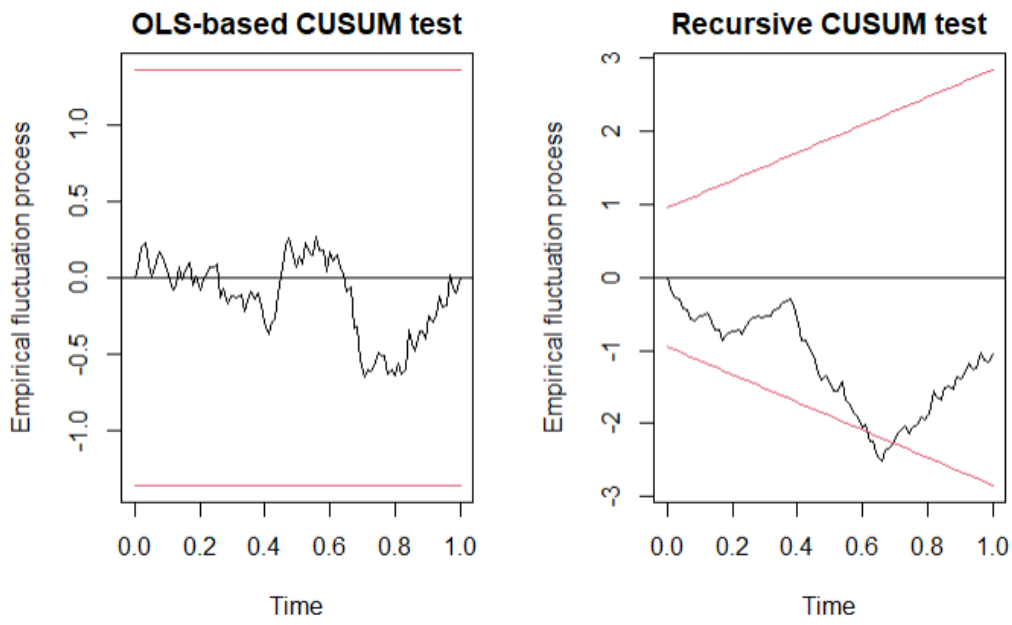
Tabel A.3: Test for heteroskedasticitet ved hjælp af Breusch-Pagan test for fejlkorrektionsmodellen.

Lags	X-squared	Df	p-value
12	12.406	12	0.4136
8	8.9474	8	0.3467
4	3.2893	4	0.5106

Tabel A.4: Box Test for seriekorrelation.

Test type	p-value
Shapiro-Wilk	0.9196
Jarque Bera	0.9315

Tabel A.5: Normalitetstest af den anvendte fejlkorrektionsmodel, der kan ses af ligning 5.2.



Figur A.3: CUSUM-test af den anvendte fejlkorrektionsmodel, der kan ses af ligning 5.2.

Chisq	Df	P-værdi
7.07	1	0.008

Tabel A.6: Wald Chi-square test om $-\frac{\phi_3}{\theta_1} = 0$.

Variabel	Langsigtede estimat	P-værdi
log_gdp	0.36*	0.0001
log_u	-0.20	0.107
Variabel	**kortsigtede estimat**	**P-værdi**
Δlog_hp_{t-1}	0.614*	0.0001
Δlog_hp_{t-4}	0.346*	0.0001
Δlog_u_t	-0.034	0.057
Δlog_u_{t-1}	0.065*	0.0001
Δlog_u_{t-2}	0.043*	0.0001
Δlog_u_{t-4}	0.052*	0.003
$\Delta log_afdragsfrihed_{t-2}$	0.394*	0.003

Tabel A.7: Yderligere effekter af den overordnede fejlkorrektionsmodel (* ved 5% signifikansniveau).

Appendiks B: Den korte fejlkorrektionsmodel

	$\Delta log_hp_short_t$ ARDL 1
$\Delta log_hp_short_{t-1}$	0.417***
	(0.119)
$\Delta log_hp_short_{t-4}$	0.238*
	(0.122)
$\Delta log_gdp_short_{t-3}$	−0.245*
	(0.133)
$\Delta log_gdp_short_{t-4}$	−0.342***
	(0.118)
$\Delta log_u_short_t$	−0.065***
	(0.020)
$\Delta log_u_short_{t-1}$	0.040**
	(0.017)
$\Delta log_u_short_{t-2}$	0.034*
	(0.018)
$\Delta afdragsfrihed_short_{t-4}$	−0.514**
	(0.210)
log_hp_short_level	−0.073**
	(0.031)
log_gdp_short_level	0.031**
	(0.012)
log_u_short_level	0.010
	(0.013)
afdragsfrihed_short_level	−0.146***
	(0.039)
N	56
Adjusted R^2	0.827

Noter: ***Signifikant ved et signifikansniveau på 1%.
**Signifikant ved et signifikansniveau på 5%.
*Signifikant ved et signifikansniveau på 10%.

Tabel B.1: Estimering af den korte fejlkorrektionsmodel.

Type af test	P-værdi
Shapiro-Wilk	0.2098
Jarque Bera	0.502

Tabel B.2: Normalitetstest af den anvendte lineære regressionsmodel, der kan ses af ligning 5.5.

BP	Df	P-værdi
3.3468	11	0.9853

Tabel B.3: Test for heteroskedasticitet ved hjælp af Breusch-Pagan test for den korte fejlkorrektionsmodel.

RESET	Df1	Df2	P-værdi
1.6845	2	42	0.1978

Tabel B.4: RESET-test for test af lineære parametre for short ecm model.

Lags	X-squared	Df	P-værdi
12	19.697	12	0.07303
8	11.659	8	0.1671
4	2.6764	4	0.6133

Tabel B.5: Box Test af den korte fejlkorrektionsmodel for seriekorrelation.

Chisq	Df	P-værdi
4.012	1	0.045

Tabel B.6: Wald Chi-square test om $-\frac{\phi_3}{\theta_1} = 0$.

Variabel	Langsigtede estimat	P-værdi
log_gdp	0.421*	0.0001
log_u	0.143	0.535
Variabel	**Kortsigtede estimat**	**P-værdi**
Δlog_hp_{t-1}	0.417*	0.001
Δlog_hp_{t-4}	0.236	0.056
Δlog_gdp_{t-3}	-0.245	0.073
Δlog_gdp_{t-4}	-0.342*	0.006
Δlog_u_t	-0.065*	0.002
Δlog_u_{t-1}	0.040*	0.026
Δlog_u_{t-2}	0.034	0.064
$\Delta log_afdragsfrihed_{t-4}$	-0.514*	0.019

Tabel B.7: Yderligere effekter af den korte fejlkorrektionsmodel (* ved 5% signifikansniveau).

Appendiks C: OLS den cykliske del af variable

Type af test	P-værdi
Shapiro-Wilk	0.083
Jarque Bera	0.098

Tabel C.1: Normalitetstest af den anvendte lineære regressionsmodel, der kan ses af tabel 16.

BP	Df	P-værdi
7.4857	6	0.2783

Tabel C.2: Test for heteroskedasticitet ved hjælp af Breusch-Pagan test for OLS modellen.

RESET	Df1	Df2	P-værdi
0.889	2	48	0.4176

Tabel C.3: RESET-test for test af lineære parametre af OLS modellen.

Lags	X-squared	Df	P-værdi
12	19.126	12	0.0855
8	16.006	8	0.0422
4	7.1335	4	0.129

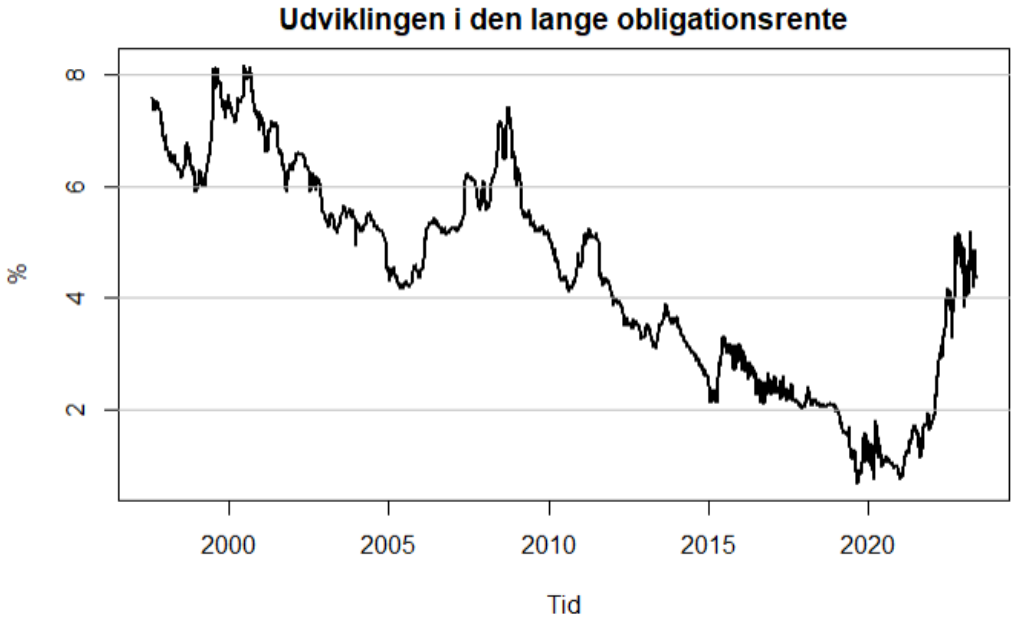
Tabel C.4: Box Test for seriekorrelation af OLS modellen.

Appendiks D: Robusthedstjek af DiD estimater

Koefficienter	Estimat	Std. Fejl	t værdi	Pr(>|t|)
tid	9.60953	5.96892	1.6099	0.121054
did	13.78710	5.25623	2.6230	0.015205 *
arbejdsløshed	-2.63872	0.71745	-3.6779	0.001247 **
rentesats	1.30811	1.70292	0.7682	0.450210
BNP-indeks	0.96080	0.31787	3.0226	0.006060 **
Signifikans koder:				
0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' ' 1

Tabel D.1: Estimater for robusthedstjek af difference in difference analysen med kontrolvariable.

Appendiks E: Andet



Figur E.1: Den gennemsnitlige lange obligationsrente (Finans Danmark, 2023).

$$
\begin{gathered}
y_t = \tau_t + c_t \\
\min_{\{\tau_t\}} \left\{ \sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T_1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right\}
\end{gathered}
\tag{8.1}
$$