



REGIONAL VÆKST

Et studie af viden og dens interregionale spillover-effekt på væksten i de europæiske regioner

Elvira Barzegar & Majbritt Jensen

Aalborg Universitet

Cand.oecon

07.08.2017

ABSTRACT

A significant amount of growth literature has highlighted the importance of knowledge in fostering and retaining growth. The concept of knowledge is however very abstract, and there are disagreements as to what qualities knowledge possess. Knowledge is famously seen as a public good by Romer (1990), and by having this quality he assess that generation of new knowledge will automatically benefit other agents in the economy, assuming it will be accessible to all in an instant. However this line of thought has been challenged by quite a few regional economic scientists. They claim that knowledge having a tacit component will not be useful to all, because it will be more difficult to communicate especially at greater geographic distances. By merely making a visual inspection of European regions, this study found large differences between European regions in terms of economic development, growth and the level of knowledge production as proxied by the number of patents (codified knowledge) and number of employees working in science and technology (tacit knowledge). Empirically there seems to be a pattern in which strong and knowledge intensive regions are geographically located close to each other whereas the economically weaker regions are concentrated in a geographically confined space, which seems to indicate that there is some form of spatial dependence between the regions. Through these observations and through trying to settle between the former mentioned theoretical dispute, this thesis attempts to model regional growth and estimate the impact of the tacit and codified component of knowledge on growth. This is done by the selection of two different albeit quite similar model specifications known as Spatial Durbin and Spatial Durbin Error Model with which it is possible to model spatial dependence. In line of this choice this thesis estimates the impact of interregional knowledge spillovers on growth assuming that knowledge does have a positive impact, however only at shorter geographic distances. In modelling the regional growth this thesis has in line with earlier work found that the benefits of knowledge seem to decay with distance and especially tacit knowledge is prone strong decays. Surprisingly tacit knowledge did not have a significant impact on regional growth, although it did not have any significant spillover effect as expected. Another interesting result of this thesis is that knowledge spillovers only have a significant impact on regional growth when considering within country regions which seems to indicate that country borders serve as a barrier for receiving the impacts of knowledge spillovers.

Aalborg Universitet
7. august 2017

Samfundsøkonomi, Cand.Oecon.
10. semester – specialeafhandling

Elvira Barzegar
Studie nr. 20123609

Majbritt Jensen
Studie nr. 20123603

Vejleder:
Ina Drejer

Anslag: 164.448
Normalsider: 68,52

Elvira Barzegar

Majbritt Jensen

Data og programmeringskoder er vedlagt på medfølgende CD. Bemærk, at data for BNP per indbygger er siden specialets udarbejdelse blevet opdateret i Eurostats database, hvorfor der kan forekomme afvigelser i resultater ved reproduktion.

Indholdsfortegnelse

1	Indledning.....	1
1.1	Motivation	1
1.2	Problemfelt	3
1.3	Projektets opbygning	4
2	Teori	5
2.1	Romer modellen	5
2.1.1	Opsummering	9
2.2	Regional vækst og interregionale spillovers af viden.....	10
2.2.1	Typer af viden.....	10
2.2.2	Diffusion af viden og geografisk rækkevidde	12
2.2.3	IKT og Death of Distance? En ekskurs	14
2.2.4	Opsummering	15
3	Eksisterende forskning	16
3.1	Tidligere forskning	16
3.2	Eget bidrag.....	19
4	Metode.....	20
4.1	Datastruktur og den ikke-rumlige model.....	20
4.2	Rumlig afhængighed	21
4.3	Rumlig økonometri.....	25
4.3.1	Modelspecifikationer	25
4.3.2	Modelestimation og fortolkning	29
4.3.3	Specifikationstest.....	33
4.4	Databeskrivelse.....	33
4.5	Den økonometriske model.....	34
4.5.1	Variabelbeskrivelse	36
4.5.2	Variable til robusthedstjek.....	47
4.5.3	Alternative variable	50
5	Resultater.....	52
5.1	Test for rumlig afhængighed	52
5.2	Estimationsresultater	53
5.2.1	Intra- og internationale spillover-effekter mellem regioner?.....	57
5.3	Robusthedstjek	58
5.3.1	Delstikprøve baseret på udviklingsniveau	59

6	Diskussion	64
6.1	64
6.2	Overnational videns- og vækstpolicy	68
7	Konklusion	72
8	Perspektivering	74
9	Litteraturliste	76
Bilag	80

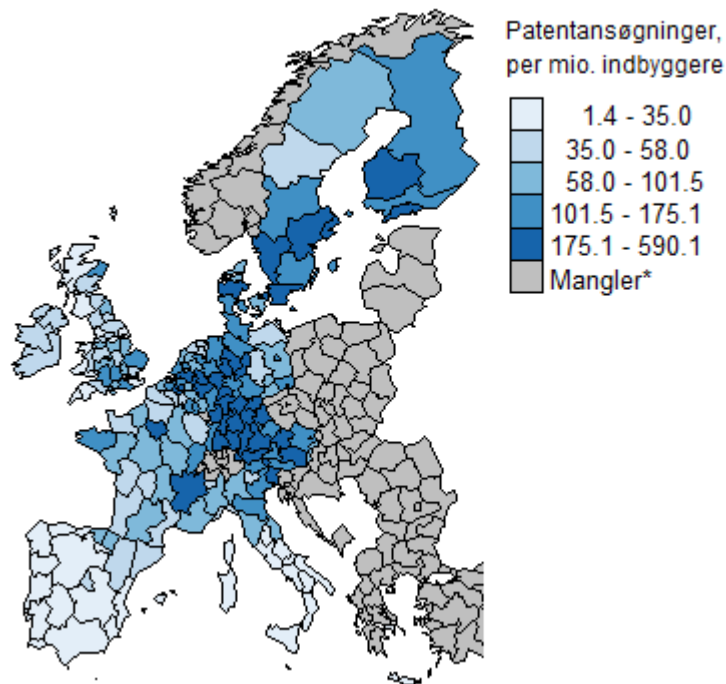
1 Indledning

1.1 Motivation

Viden har længe været anerkendt som en vigtig ressource for økonomien. I starten af 1900-tallet havde flere prominente økonomer betonet vigtigheden af viden for funktionen af flere aspekter i økonomien, herunder for produktion, organisation og markeder (Cooke & Leydesdorff, 2006, s. 7). Heriblandt havde Schumpeter genkendt vigtigheden af nye kombinationer af viden i relation til innovation og entreprenørskab, mens Marshall beskrev hvordan viden udgør en vigtig del af kapital, og er derved en stærk motor bag produktion (Cooke & Leydesdorff, 2006, s. 6). Det er dog først i nyere tid at viden er blevet modelleret eksplicit ind i teorier og modeller med eksempelvis *New Growth Theory* fra 1980'erne, hvor viden blev inkorporeret i makroøkonomisk vækstteori med henblik på at forklare hvordan viden øger produktivitet og økonomisk vækst (fx Romer, 1990).

Det øgede fokus på viden i den teoretiske verden er kommet i takt med, at output i vidensintensive og højteknologiske erhverv eksploderede (OECD, 1996, s. 9). Med dette menes at virksomheder bruger mere end 4 pct. af deres omsætning på investeringer i forskning og udvikling (Smith, 2000). I 1990'erne nåede vidensbaseret output at udgøre mere end 50 pct. af BNP i de store OECD-økonomier, og denne udvikling resulterede i begrebet "*knowledge-based economy*", som blev brugt til at karakterisere førnævnte økonomier (OECD, 1996, s. 9). Eksempler på vidensbaseret output inkluderer computere, medicinalvarer og luftfarttøjer, og disse produceres typisk i højteknologiske industrier, som er vidensintensive (OECD, 1996, s. 36). Vidensbaserede økonomier er kort sagt kendetegnet ved, at de i høj grad er baseret på vidensproduktion, vidensdistribution og udnyttelse af viden, hvilket er i modsætning til økonomier, som hovedsageligt er afhængige af naturressourcer (OECD, 1996, s. 7). Hvad der dog er evident ud fra en visuel inspektion af Figur 1 er, at produktionen af viden, her målt på antal patentansøgninger, er geografisk koncentreret indenfor forskellige områder af de forskellige lande.

Figur 1 Antal ansøgte patenter per mio. indbyggere i år 2012 i EU-15 lande.



Note: * Markerer enten at data mangler, eller at regionerne er udenfor EU-15.
 Kilde: Eurostat (2017a), © EuroGeographics for de administrative grænser.

Denne observation inviterer til spørgsmålet: Hvorfor er den geografiske fordeling af vidensproduktion relevant i en økonomisk sammenhæng? I konteksten af vækst har flere økonomer identificeret den geografiske koncentration af vidensproduktion, som en mulig forklaring på den ujævne geografiske fordeling af vækst blandt landes regioner (Balland & Rigby, 2017, s. 2). Som det fremgår af Tabel 1 er spredningen i vækstperformance endda større blandt regioner end den er iblandt lande.

Tabel 1 Spredning i vækst i BNP per indbygger på tværs af EU-15 lande og NUTS 2 regioner, 2002-2014

Δ BNP per indbygger		
Lande	Min.	0,212 pct. (Grækenland)
	Max.	3,038 pct. (Luxembourg)
	Spredning	2,826 pp.
Regioner	Min.	-1,2775 pct. (Sterea Ellada)
	Max.	3,96573 pct. (Braunschweig)
	Spredning	5.25105 pp.

Kilde: Egne beregninger på baggrund af data fra Eurostat (2017a).
 Baseret på (OECD, 2009a, s. 23)

Der er en grund til at være opmærksom på regioner der har et lavere udviklingsniveau, målt på BNP per indbygger, idet de ifølge OECD har stået for mere end 50 pct. af væksten i BNP i deres respektive lande i perioden 1995-2005 (OECD, 2009b, s. 14). Eksempelvis bidrog de regioner i Østrig og Sverige, hvis BNP lå under gennemsnittet, med 53 pct. af den nationale vækst (OECD, 2009b, s. 33). Når disse regioner står for så stor en del af væksten i deres lande, anbefaler OECD at de tilbagestående regioner også udnytter deres potentialer. Målet er ikke at de skal bestræbe sig på at tilnærme hinandens BNP-niveau, men at regionerne finder og udnytter potentielle vækstkilder, for at øge deres velstand (OECD, 2009b, s. 27).

Selvom regioner selv kan igangsætte initiativer for at fremme produktionen af viden, så er der en indikation på, at naboregioners vidensproduktion også har betydning for egen vidensproduktion, idet der jf. Tabel 1 ser ud til at regioner med en høj produktion af viden er geografisk tæt placerede. Således er det nærliggende at undersøge om naboregioners vidensproduktion, foruden egen vidensproduktion, også har betydning for regioners vækst.

1.2 Problemfelt

Med udgangspunkt i ovenstående overvejelser, søger dette speciale at besvare følgende problemformulering:

”Hvilken effekt har viden og interregionale vidensspillovers på regioners vækstperformance?”

Denne problemformulering besvares kvantitativt ved at tage udgangspunkt i data for de europæiske regioner i EU-15. For at undersøge relationen mellem en regions viden og vækst empirisk, så kræver det at viden kan kvantificeres. Viden er dog en svær størrelse at måle, da det er et bredt begreb. Ved hjælp af økonomisk teori kan viden disaggregeres i forskellige typer, så det kan blive lettere at opfange forskellige aspekter af viden, og dermed have mulighed for at kvantificere viden. Ved at afdække relationen mellem viden og regional vækst, så kan der findes indikationer på, hvorvidt regioner rent policymæssigt skal fokusere deres indsatser på vidensintensive aktiviteter. Derudover er der en tendens til at andre regioners vidensproduktion også har en betydning for væksten i en region, altså om der sker en spillover af viden. Ved at inkludere spillover-effekter, så tages der også højde for at regional vækst kan påvirkes af eksterne forhold udover egen indsats.

1.3 Projektets opbygning

Her præsenteres, hvordan specialet er opbygget og indholdet af hvert kapitel.

I *kapitel 2* fremlægges specialets teoretiske referenceramme, hvor udgangspunktet er Romers (1990) endogene vækstteori, og dernæst introduceres regional udviklingslitteratur, der beskriver hvilken rolle forskellige komponenter af viden og geografi spiller for rækkevidden af spillover-effekter af viden.

I *kapitel 3* præsenteres eksisterende forskning på området, for at kunne skabe en kontekst for dette speciales genstandsfelt og metode.

Kapitel 4 udgør metodeafsnittet, hvor den økonometriske metode præsenteres. Indledningsvist præsenteres konceptet rumlig afhængighed, og dernæst præsenteres den anvendte litteratur omkring rumlig økonometri. Herefter beskrives den indsamlede data som er brugt til at estimere modellerne. Efterfølgende opstilles den økonometriske model på baggrund af den præsenterede teori og økonometri litteraturen, efterfulgt af en beskrivelse af de inkluderede variable.

I *kapitel 5* præsenteres resultaterne fra regressionen af den økonometriske model. Derudover testes robustheden af resultaterne overfor modelspecifikationerne.

I *kapitel 6* fortolkes og diskuteres resultaternes implikationer.

2 Teori

Formålet med nærværende afsnit er at præsentere teoretiske argumenter for hvad der skaber en langsigtet økonomisk vækstrate i output per indbygger i regioner. Derfor præsenteres indledningsvist Romers endogene vækstmodel med henblik på at afdække, hvad der kan skabe økonomisk vækst. Romers indsigt suppleres dernæst med regional udviklingslitteratur, som bl.a. betoner betydningen af geografisk nærhed og videns egenskaber, samt konsekvensen af disse for spillover af viden.

2.1 Romer modellen

Det teoretiske afsæt der tages i Romer (1990) skyldes at Romer heri endogeniserer teknologisk udvikling, og at netop teknologisk udvikling er drivkraften for et lands langsigtede økonomiske vækstrate. Modellen stammer fra hans berømte artikel "Endogenous Technological Change" fra 1990, og er et af hovedbidragene indenfor endogen vækstteori. Romers model adskiller sig fra vækstteorier af ældre dato, idet teknologisk (videns) udvikling indtil da var eksogent givet, se f.eks. (Solow, 1956), eller blot ansås for at være et tilfældigt biprodukt af *learning-by-doing*, se f.eks. (Lucas, 1988; Arrow, 1962). Uanset kilden til teknologisk udvikling, såfremt der ligger viden bag, så er der i økonomisk litteratur overvejende enighed om, at teknologisk udvikling skaber langsigtet vækst i output per indbygger. Ifølge Romer defineres teknologisk udvikling som:

"[...] improvement in the instructions for mixing together raw materials". (Romer, 1990, s. 72)

Teknologisk udvikling er drevet af bevidste investeringer i forskning og udvikling (F&U) foretaget af profit-maksimerende agenter, derved åbnes der op for at et land selv kan være en aktiv spiller i at øge den teknologiske udvikling, og derved være med til at øge væksten i landet. Den teknologiske udvikling i modellen bygger på akkumulering af viden og skabelse af ny viden. Romer skelner mellem to typer af viden, herunder viden bundet til humankapital og viden bundet til designs. Den type af viden, der ifølge Romer driver væksten, er den viden der er bundet til teknologi, A , og som Romer kalder for *designs* (Romer, 1990, s. 79). Når omkostningerne til teknologi er afholdt, så kan den viden der er bundet til A bruges igen og igen uden yderligere omkostninger (Romer, 1990, s. 72).

Viden bundet til humankapital, H , tillægges personerne i arbejdsstyrken, og dækker over en persons uddannelsesniveau og erhvervs erfaring, hvorfor den akkumulerede viden ikke er ens hos alle, hvilket gør H til et *rivaliserende* input. Det vil sige at når en person ikke længere er en del af arbejdsstyrken, går den viden de har akkumuleret tabt (Romer, 1990, s. 75). Teknologi er derimod ikke bundet til en person, og er noget alle har adgang til, hvorfor den er et *ikke-rivaliserende* input.

Dette betyder også at A ikke har nogen øvre grænse, dvs. der er ikke en øvre grænse for hvor mange designs der kan opfindes. Samtidig kan det enkelte design kendetegnes ved at være *delvis ekskluderende* (Romer, 1990, s. 75). Det vil sige, at ejeren af et design har rettigheden til dets brug i produktion, og kan ekskludere andre fra at bruge designet i denne sammenhæng, men ejeren kan ikke ekskludere forskere fra at bruge den viden, der er bundet til designet, hvorfor design kun er delvis ekskluderende (Romer, 1990, s. 84). Et design må ifølge Romer have denne egenskab for at give agenter et incitament til at udvikle ny viden, og gør det muligt at afholde *free riders* fra at udnytte andres designs, ved at tildele dem patenter og derved ekskludere dem.¹ Dette betyder at der i denne model åbnes op for ufuldkommen konkurrence (Romer, 1990, s. 81).

Idet hver ny enhed af viden (i relation til teknologi) ifølge Romer korresponderer med et design for et nyt gode (Romer, 1990, s. 79), kan teknologisk udvikling i Romers model tolkes som værende svarende til udvikling af viden.

Modelantagelser

Modellen i Romer (1990) har fire inputs i produktionsprocessen herunder kapital (K), arbejdskraft (L), humankapital (H) og det teknologiske niveau (A) (Romer, 1990, s. 78). K måles i antallet af forbrugsgoder. L er antallet af mennesker i arbejdsstyrken. H er den akkumulerede viden en ansat har opnået gennem tiden i form af uddannelse og erhvervs erfaring, og er unik fra person til person. Derudover består modellen af tre sektorer, forskningssektoren, kapitalgodesektoren og færdigvaresektoren (Romer, 1990, s. 79).

Forskningssektoren gør brug af humankapital og det aktuelle teknologiniveau til bl.a. at producere ny viden, men især nye designs af varige forbrugsgoder. Når ny viden produceres vil den altid få en kommerciel værdi, og vil altid kunne afsættes til en virksomhed i kapitalgodesektoren.

Kapitalgodesektoren bruger de designs som forskningssektoren opfinder, for sammen med de i forvejen tilgængelige designs, at producere de kapitalgoder som færdigvaresektoren skal bruge til at producere det endelige output med. Et kapitalgode kan kun produceres af én virksomhed.

Virksomheder i kapitalgodesektoren køber eller fremstiller selv et design før det må produceres, og derefter har virksomheden et evighedspatent på designet. Når en virksomhed ejer rettighederne til godet kan de omdanne η enheder af inputs til kapitalgodet i . Herefter lejer virksomheden godet ud til virksomheder i færdigvaresektoren til prisen $p(i)$. Dette betyder at virksomheden der ejer

¹ En free rider er en person, som benytter sig af andres goder, uden at yde en økonomisk kompensation herfor.

patentet har monopol på netop dette design, og derfor er der i kapitalgodesektoren ufuldkommen konkurrence (Romer, 1990, s. 82).

Færdigvaresektoren bruger arbejdsstyrken, humankapital og de tilgængelige kapitalgoder til at producere det endelige output. I færdigvare sektoren er der konstant skalaafkast, dette betyder at output i færdigvaresektoren kan beskrives som værende produktet af en enkelt pristagende virksomhed. Dette betyder at der i færdigvaresektoren er fuldkommen konkurrence og virksomheden kan ikke påvirke prisen ved at ændre i den mængde virksomheden producerer.

Modellen indeholder ifølge Romer selv en række simplificerende antagelser (Romer, 1990, s. 80). En af disse antagelser er, at populationen og arbejdsstyrken skal være konstant. Derved undgår Romer at faktorer som arbejdsmarkedsdeltagelse og fertilitet påvirker væksten. En anden antagelse er, at den totale beholdning af humankapital i befolkningen er holdt fast, og det samme er den del der er til rådighed for markedet. Det betyder at den totale beholdning af humankapital ikke kan udbygges. En mere ekstrem antagelse er, at kapitalakkumulation kan ske ved at flytte ressourcer fra færdigvaresektoren over til kapitalgodesektoren, og at disse er separate sektorer, men med adgang til samme teknologi. Derudover antages det, at det kun er viden og humankapital der indgår i produktionen af nye designs og ny viden. Arbejdsstyrken og kapital indgår således ikke (Romer, 1990 s. 80).

Modellen

Ud fra ovenstående opstiller Romer en model, hvor output, Y , er en funktion af arbejdskraft, L , og den del af humankapitalen der bruges i færdigvaresektoren, H_Y , og fysisk kapital, x (Romer, 1990, s. 80). Produktionsfunktionen i denne model tager form af en Cobb-Douglas produktionsfunktion og kan derfor skrives som følgende:

$$Y(H_Y, L, x) = H_Y^\alpha L^\beta \sum_{i=1}^{\infty} x_i^{1-\alpha-\beta} \quad (1)$$

H_Y er den del af humankapitalen, der anvendes i færdigvaresektoren. På det aggregerede niveau er der en sammenhæng mellem H_Y og H_A , fordi de skal opfylde betingelsen: $H_Y + H_A = H$. Det betyder at enhver person kan dedikere sin humankapital til enten forskningssektoren eller til færdigvaresektoren. Her ses der dog bort fra at humankapital og arbejdskraft tilføres sammen. Helt bogstaveligt betyder det at personer der specialiserer sig i humankapitalakkumulation ikke bidrager med sin arbejdskraft (Romer, 1990, s. 85).

For at undgå at der opstår et problem med *free riders* i relation til forskningssektorens output, så introduceres patenter, som et incitament for forskere til at investere deres humankapital i produktionen af ny viden. Patentet på den nye viden bliver derefter solgt til en virksomhed i kapitalgodesektoren til prisen P_A . Når en virksomhed vælger at investere i rettighederne til at producere et nyt kapitalgode afhænger det af den forventede fortjeneste i forhold til investeringsomkostningerne (Romer, 1990, s. 87). Da markedet for nye designs er konkurrencepræget, vil prisen på et nyt design blive budt op, og lande på en pris der svarer til nutidsværdien af den samlede profit en monopolist vil kunne sikre sig (Romer, 1990, s. 87). Prisen på et patent kan derfor udtrykkes således:

$$\pi(t) = r(t)P_A \quad (2)$$

Ifølge denne ligning skal fortjenesten til enhver tid kunne dække renteudgifter på investeringen i det nye design (Romer, 1990, s. 87).

Teknologisk udvikling

I og med at viden i form af designs er et ikke-rivaliserende gode, så vil viden kunne udbredes til andre agenter, dvs. der sker såkaldte spillovers af viden. Grundet spillovers, så kan A også betragtes som en vidensbeholdning, som alle forskere til enhver tid har adgang til og kan benytte sig af. Hvis en forsker er i besiddelse af en vis mængde humankapital, H^j , så vil den mængde af nye designs produceret af forsker, j , kunne udtrykkes som $\delta H^j A$, hvor δ er en produktivitetsparameter (Romer, 1990, s. 83). Det betyder at udviklingen i den aggregerede sum af nye designs produceret af hele forskningssektoren udtrykkes som følgende:

$$\dot{A} = \delta H_A A \quad (3)$$

Udtrykket indeholder følgende antagelser (Romer, 1990, s. 83):

- 1) Jo mere humankapital der dedikeres til forskning og udvikling, jo mere stiger teknologiniveauet.
- 2) Jo større den totale beholdning af viden og designs er, jo højere vil produktiviteten hos forskerne være.

Modellens løsning

For at finde modellens løsning kræver det at der findes et ligevægtsniveau, hvor variablerne A, K og Y har en konstant eksponentiel vækstrate (Romer, 1990, s. 90). Intuitionen er at en ligevægtssituation vil kunne finde sted hvis A vokser med en konstant eksponentiel vækstrate. Dette kan lade sig gøre fordi \dot{A} i ligning 3 er lineær i A . For at dette er opfyldt, skal den mængde af humankapital der er dedikeret til forskningssektoren holdes konstant. Dette betyder at der kan eksistere en ligevægtssituation og derfor skal det sikres at priser og lønninger ligger på et niveau der gør at H_Y og H_A forbliver konstante når Y, K, C og A vokser. Modellen antager at outputtet vokser med samme rate som A hvis L, H_Y og \bar{x} holdes fast (Romer, 1990, s. 90). \bar{x} er niveauet kapitalgoderne udbydes på, hvilket derved betyder at alle kapitalgoder skal udbydes på samme niveau. Hvis \bar{x} er fast så må K vokse med samme rate som A , fordi den totale brug af kapital er $A\bar{x}\eta$. Da forholdet mellem K og Y er konstant må forholdet:

$$\frac{C}{Y} = 1 - \frac{\dot{K}}{Y} = 1 - \frac{\dot{K} K}{K Y} \quad (4)$$

også være det (Romer, 1990, s. 92). Hvis g er vækstraten for A, Y og K vil væksten i output kunne udtrykkes som:

$$g = \frac{\dot{C}}{C} = \frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} = \frac{\dot{A}}{A} = \delta H_A \quad (5)$$

Det vil betyde at væksten i outputtet er afhængig af produktiviteten i forskningssektoren, og derfor hvor meget ny viden der produceres (Romer, 1990, s. 92). Dette betyder også at g ikke kan være negativ.

2.1.1 Opsummering

Ved at endogenisere teknologisk udvikling har modellen åbnet op for, at investering i ny viden er en bevidst handling, og at et land selv har mulighed for at påvirke væksten i sin økonomi. En af modellens vigtigste bidrag i forhold til projektet er derfor, at jo større økonomiens samlede beholdning af humankapital er, desto større vækstrater vil de opleve, fordi det betyder at der kan udvikles flere designs, som er den type viden, der ifølge Romer driver væksten.

2.2 Regional vækst og interregionale spillovers af viden

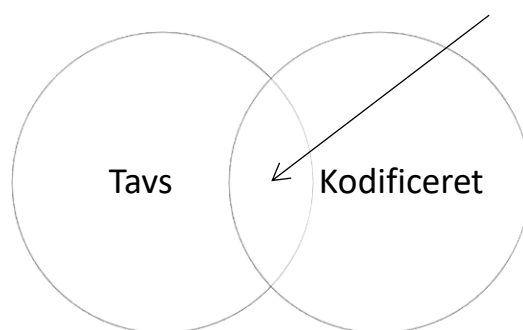
Analysegenstanden for Romers vækstteori er en national økonomi, og økonomien er lukket, dvs. der tages ikke højde for udefrakommende påvirkninger. Vækst observeret på det makroøkonomiske niveau er hos Romer (1990) et resultat af de økonomiske agents adfærd og samspil på det mikroøkonomiske niveau, men implikationen af en lukket økonomi er, at spillover af viden kun sker mellem agenter indenfor økonomien og ikke på tværs af økonomier. I regional udviklingslitteratur anses spillovers af viden dog for at være et geografisk afgrænset fænomen, så i dette henseende vil viden, der bliver skabt i en bestemt geografisk beliggenhed i et land ikke nødvendigvis udbredes og anvendes i alle andre beliggenheder i landet. Selvom spillovers af viden er et geografisk afgrænset fænomen, så er der ifølge Audretsch & Feldman (2004, s. 6) ikke grund til at tro, at spillovers af viden bremses af by-, stats- eller landegrænser. Der er således en god grund til at operere med en analyseenhed, der er mindre end en nationaløkonomi, såsom en regionaløkonomi, og ligeledes ikke se på den i isolation, men derimod i relation til andre regioners økonomier indenfor og på tværs af landegrænser. Ved at operere med regioner fremfor eksempelvis nationer, så kan den *rumlige dimension* også spille en aktiv rolle i generering af vækst. Regional udviklingslitteratur fremhæver rummet, som værende en ressource i sig selv, idet rummet kan være kilde til stigende skalafkast og positive territoriale eksternaliteter i form af agglomeration (Capello, Space and Theoretical Approaches to Regional Growth, 2008, s. 22). Hvordan dette gør sig gældende, og i højere grad hvad det indebærer for spillovers af viden, vil være fokuset i følgende afsnit. Først vil der dog redegøres for hvilken anskuelse af begrebet viden, der opereres med i nærværende studie.

2.2.1 Typer af viden

Viden optræder typisk i *tavs* og/eller *kodificeret* form (*tacit* og *codified*). Kodificeret viden er en form for eksplicit viden, som groft sagt kan oversættes til ”information” (Foray & Lundvall, 1998, s. 117; Cowan, David, & Foray, 2000, s. 218). Denne form for viden kan eksempelvis findes i form af bøger eller patenter, og denne form for viden er typisk lettere at overføre mellem mennesker i modsætning til tavs viden. Begrebet tavs viden kan spores tilbage til Michael Polanyi, som konstaterede at mennesker kan ”vide mere end de kan fortælle” (Polanyi, 1966, s. 4). Det vil sige, at mennesker besidder en form for viden, som i vid udtrækning ikke kan verbaliseres eller dokumenteres. Denne egenskab gør tavs viden relativt sværere at overføre mellem mennesker. Et eksempel på viden, som generelt findes i tavs form er, hvad Lundvall & Johnson (1994) kalder for *know-how*, som er en form for viden der dækker over opnåede færdigheder, som er svær at

kodificere, idet den opnås igennem personlig erfaring og learning-by-doing.² Hvorvidt en portion viden er kodificeret eller tavs, afhænger også af den tidsmæssige, rumlige, kulturelle og sociale kontekst (Cowan, David, & Foray, 2000, s. 225). Der kan som regel ikke antages, at kodificeret viden er standardiseret eller med andre ord universel. Hvad der er kodificeret for en person eller gruppe, kan være tavs for andre, og kan endda være et mysterium for en tredje person eller gruppe (Cowan, David, & Foray, 2000, s. 225). Dette aspekt er ifølge Cowan et al. (2000, s. 226) problematiserende for den endogene vækstteoris antagelse om en ”verdensbeholdning af viden”, som formodningsvist må bygge på, at alt der kodificeres bliver frit tilgængelig, og dermed deler karakteristika med et offentligt gode. I stedet ville en anden definition af ”verdensbeholdning af viden” være mere passende. Ifølge Cowan et al. (2000, s. 226) skulle den i stedet defineres, som en fællesmængde af viden, der består af viden der *deles* mellem agenter. I og med at meget viden ikke deles, så må denne fællesmængde af viden være meget lille (Cowan, David, & Foray, 2000, s. 226). Nyttigheden af kodificeret viden beror således på, at den rent faktisk er tilgængelig, og at der haves den fornødne tavse viden, til at kunne anvende den kodificerede viden. Således er der implicit hos Cowan et al. (2000) en forståelse af, at tavs viden og kodificeret viden ikke nødvendigvis udgør to urelaterede størrelser, men at de komplementerer hinanden, og at brugen af den ene kan forudsætte at have den anden, som illustreret i Figur 2.

Figur 2 Venn diagram over relationen mellem tavs og kodificeret viden i brugen af viden.



Kilde: Egen tilvirkning, baseret på Cowan et al. (2000).

² Lundvall & Johnson (1994) opererer med en typologi af (økonomisk) viden, hvor der skelnes mellem fire typer af viden, som udover *know-how* inkluderer *know-what*, *know-why* og *know-who*, se (Lundvall & Johnson, 1994, s. 27-28). Disse typer af viden er i forskellige grader tavs eller kodificeret, alt efter kontekst.

Figur 2 viser hvorledes viden består af en tavs komponent og en kodificeret komponent, men hvor der ikke eksisterer klare skillelinjer mellem hvilken type af viden der bruges, fordi brugen af kodificeret viden kan være afhængig af en relevant tavs viden.

2.2.2 Diffusion af viden og geografisk rækkevidde

Den blotte eksistens af viden sikrer ikke, at den automatisk bliver spredt ud og udnyttet af alle agenter. Distinktionen mellem tavs og kodificeret viden har nemlig implikationer for hvordan spillovers af viden forstås i et geografisk rum. Når der eksisterer en geografisk koncentration (agglomeration) af økonomiske agenter, så vil rummet danne baggrund for udveksling af viden, og især tavs viden, på grund af *geografisk nærhed*.³ Geografisk nærhed refererer til en nærhed mellem økonomiske agenter i rumlig eller fysisk forstand, og anses i regional udviklingslitteratur som værende fremmende for spillovers af viden (Boschma, 2005, s. 69). Nærhed fremmer udbredelsen af tavs viden, idet den typisk spredes gennem ansigt-til-ansigt interaktioner (Storper & Venables, 2004). Således har tavs viden en relativ kort geografisk rækkevidde, hvorimod kodificeret viden har en potentiel større geografisk rækkevidde, fordi den i en mindre grad er afhængig af ansigt-til-ansigt interaktion, men derimod kan udbredes næsten omkostningsfrit gennem informationsstrukturer (Leamer & Storper, 2001, s. 650). Geografisk nærhed giver således anledning til agglomerationsfordele, fordi økonomiske agenter opnår en gevinst ved at kunne placere sig tæt på andre agenter, idet de kan udnytte den viden som de andre agenter har akkumuleret, til at producere ny viden eller goder. Dette udelukker dog ikke, at der kan opstå ulemper ved at være i en agglomeration. Når befolkningstætheden stiger, stiger presset på områdets infrastruktur, og det kan derved bl.a. både tage længere tid og bliver dyrere for pendlere at komme frem og tilbage mellem bopæl og arbejdsplads. Andre ulemper er stigende boligpriser. Dette kan antages af have en negativ effekt på produktionen, og herved også produktionen af viden (Sedgley & Elmslie, 2004, s. 112).

Der findes adskillige eksempler på mekanismer, som understøtter spillovers på det mikroøkonomiske niveau. Disse inkluderer uddannelse, IKT, F&U-samarbejde mellem universiteter og virksomheder, og ikke mindst mobilitet af folk med særlig viden og færdigheder (Karlsson, Warda, & Gråsjö, Knowledge Spillovers in Europe: A Meta-Analysis, 2012, s. 8). Selvom der er en begrænset geografisk rækkevidde for spillovers af viden, så kan mekanismerne i henhold til Audretsch & Feldman (2004) også virke på tværs af geografiske områders grænser. I en rumlig

³ Geografisk nærhed måles typisk i fysisk distance, f.eks. antal kilometer, som vil være tilfældet i dette projekt, men alternative mål inkluderer også rejsetid i antal minutter/timer, som f.eks. hos (Rodriguez-Pose & Crescenzi, 2008, s. 58).

forstand kan der således skelnes mellem *intraregional* og *interregional* spillovers af viden (Karlsson & Gråsjö, 2012). En intraregional spillover omfatter spillover af viden mellem økonomiske agenter indenfor en region, mens sidstnævnte omfatter spillovers mellem agenter, der er på tværs af regioner.⁴

Geografisk nærhed fremmer således spillover af viden, men der kan være andre faktorer, der svækker effekten af disse, selvom der er geografisk nærhed. Flere økonomiske geografer betoner betydningen af andre nærhedsbegreber, idet geografisk nærhed ikke nødvendigvis skaber en gnidningsfri spillover af viden (Boschma, 2005; Basile, Capello, & Caragliu, 2011). Basile, Capello & Caragliu (2011, s. 25) opererer med begrebet *relationel nærhed*, som definerer hvor tætte to geografiske områder er i relation til adfærdsnormer, kultur, gensidig tillid og samarbejdsevner. Jo mere relationel nærhed der er mellem to regioner, desto lettere vil de kunne lære af hinanden, fordi der er et element af samarbejde i læringsprocesser, og desto lettere vil viden kunne flyde imellem regionerne og blive anvendt (Basile, Capello, & Caragliu, 2011, s. 24-25).⁵ I dette henseende kan den relationelle afstand mellem to regioner på tværs af landegrænser potentielt være større end regioner der befinder sig indenfor samme land, hvorfor landegrænser potentielt kan svække effekten af spillovers. I visse henseender kan relationel nærhed dog ses som værende en funktion af geografisk nærhed, da folk i tæt geografisk placering ofte – dog ikke altid – vil have fælles sprog, kultur og normer (Karlsson, Warda, & Gråsjö, 2012, s. 9).

Ydermere skal modtageren af ny viden have den relevante know-how for at kunne anvende den nye viden. Det er ikke nok at have adgang til viden, der skal med Cohen & Levinthals (1990) terminologi have en *absorptiv kapacitet*. Dette indebærer både en tilstrækkelig kognitiv kapacitet og en villighed til at tilegne sig ny viden (Döring & Schnellbach, 2006, s. 380). Modtageren skal f.eks. være villig til at pådrage sig omkostninger for at tilegne sig ny viden, hvis det er økonomisk

⁴ Der findes to traditioner, som i sin tid har prøvet at forklare intraregionale spillovers af viden. Den første tradition går tilbage til Marshall (1890), og ifølge denne vil virksomheder som tilhører samme industri have lettere ved at afkode og udnytte ny viden som skabes intraindustri end ekstraindustri (Kalapouti & Varsakelis, 2015, s. 763). Derfor vil en koncentration af virksomheder indenfor samme industri i en region promovere intraindustrielle spillovers af viden. Den anden tradition har sine aner i Jacobs (1969), og her er intuitionen i stedet, at en region med en bred vifte af industrier vil have en stor diversitet i viden, og med denne diversitet kan der bedre genereres ny viden og derved vækst (Kalapouti & Varsakelis, 2015, s. 764). Spillovers indenfor de to traditioner er også kendt som Marshal-Arrow-Romer-(MAR) spillovers og Jacobs-spillovers.

⁵ Basile, Capello & Caragliu (2011, s. 25)'s definition af relationel nærhed overlapper delvist med begrebet *social nærhed*, som beskriver graden af tillid der er mellem agenter, som interagerer med hinanden (Boschma, 2005, s. 64), og *institutionel nærhed*, som dækker over hvor tætte to regioners institutioner er, når institutioner kan forstås som et sæt af fælles kulturelle normer og vaner (uformelle institutioner), og love og regler (formelle institutioner), som regulerer folks adfærd (Boschma, 2005, s. 68).

rationelt (Döring & Schnellenbach, 2006, s. 380). Det vil kun være økonomisk rationelt hvis den kognitive distance ikke er for stor, er den det, så vil omkostningerne til "search and imitation" løbe op, og så vil det ikke kunne betale sig at forsøge med at udnytte ny viden (Boschma, 2005, s. 63). Evnen til at absorbere ny viden kræver således også kognitiv nærhed.

2.2.3 IKT og Death of Distance? En ekskurs

Umiddelbart kan det for nogen forekomme irrelevant at overveje spillover af viden i et geografisk rum med udbredelsen af moderne informations- og kommunikationsteknologi (IKT), som bryder geografiske grænser som barriere for spillover af viden. Der er blandt nogle økonomiske geografer endda blevet sået tvivl om i hvilken udstrækning distance fungerer som en hindring for udbredelsen af viden, selv med eksistensen af tavs viden (Morgan, 2004, s. 3). Blandt tvivlerne er der blevet erklæret "the death of distance" eller "the death of geography" (Feldman & Avnimelech, 2011, s. 157-158; Morgan, 2004). Kevin Morgan udfordrer til gengæld skepsissen overfor det geografiske rums betydning. Baggrunden for skeptikernes erklæring bygger ifølge Morgan (2004, s. 5) på to forhold: eksistensen af IKT, og en større vægtning af andre nærhedsbegreber end geografisk nærhed. IKT transformerer for det første måden hvorpå goder handles, og forårsager at forsyningen af services bliver frigjort fra hvor de forbruges i et geografisk rum. For det andet betragtes kodificeret viden som værende eksplicit og en universel tilgængelig form for information, som i kraft af dens egenskaber kan overføres over længere distancer både hurtigere og billigere med IKT.

Morgan er ikke uenig i, at IKT på makroniveau har transformative effekter, men han mener at skeptikerne fejlagtigt ser cyberspace som værende en reel erstatning for det geografiske rum, og at de således fejlagtigt betoner virtuel nærhed fremfor geografisk nærhed (Morgan, 2004, s. 5). Han mener, at det er svært at forestille sig at virtuel nærhed kan matche geografisk nærhed, idet sidstnævnte er rigere på diversitet, idet nuancer af kropssprog og ansigt-til-ansigt kommunikation, kan overbringe lige så meget, hvis ikke mere end verbal kommunikation alene kan (Morgan, 2004, s. 5).

Det er dog vigtigt ifølge Morgan ikke at betragte geografisk og virtuel nærhed (rum) som to parallelle mål (rum), men i stedet som to dimensioner der krydser hinanden i en kompleks facon. Således er IKT og ansigt-til-ansigt kommunikation ikke gensidigt ekskluderende, da disse sammen kan udvikle sig til at være komplementære mekanismer, f.eks. e-mails og arbejdsgange, og kan kombineres afhængig af hvilken type af transaktioner der udføres, og afhængig af fortroligheden mellem parterne (Morgan, 2004, s. 5). Ud fra dette synspunkt fremmer disse to dimensioner

hinanden, og derfor bliver det geografiske rums betydning ikke reduceret af eksistensen af det virtuelle rum.

2.2.4 Opsummering

Romers makroøkonomiske vækstteori kan bruges til at få en indsigt i hvordan regioner endogen kan skabe vækst gennem teknologisk (videns) udvikling, mens regional udviklingslitteratur supplerer med en udvidet forståelse af viden samt mikro-adfærdsmæssige indsigter, der kan bruges til at belyse hvilken rolle geografisk nærhed spiller for udbredelsen af viden.

3 Eksisterende forskning

I nærværende afsnit præsenteres tidligere empiriske studier indenfor regional vækst og interregional spillover af viden. Herunder hvordan de operationaliserer spillovers, og hvordan de håndterer rumlig afhængighed i data.

3.1 Tidligere forskning

De europæiske regioner har ofte været et genstandsfelt for studier om regional vækst. Her har studier bl.a. bestået i at undersøge effekten af forskning og udvikling (F&U), humankapital og spillovers af viden på den regionale vækst. Det er blandt disse studier dette projekt har søgt inspiration.

Regional vækst og spillover af viden er et tema i Rodríguez-Pose & Crescenzi (2008). Her undersøges effekten af vidensproduktion og spillover af viden på regional vækst i perioden 1995-2003 i de europæiske regioner i EU-25 landene, ved at opstille en økonometrisk model. Spillovers operationaliseres i denne artikel ved at konstruere en vægtmatrice, der ganges på de uafhængige variable. Vægtmatricen er en invers distancematrice, hvor distancen er defineret ud fra rejsetiden mellem region i og region j . Intuitionen med denne vægtmatrice er, at jo længere rejsetid der er mellem to regioner, jo større omkostninger vil der være forbundet med at tilegne sig andre regioners viden, hvorfor spillover vil aftage med distancen (Rodríguez-Pose & Crescenzi, 2008, s. 58).

Artiklens overordnede konklusion er, at en regions vækstpotentiale afhænger positivt af den rette sammensætning af sociale forhold kombineret med egne investeringer i F&U. Spillover af viden er samtidig vigtig, men det kræver en vis absorberingsevne, hvilket gør at den kun er signifikant ved bestemte sociale variable, nemlig uddannelsesniveaut blandt befolkningen og uddannelsesniveaut i arbejdsstyrken (Rodríguez-Pose & Crescenzi, 2008, s. 61). Desuden finder de at spillover-effekten falder med distancen. Helt præcist finder de at den forsvinder efter en rejsetid på 3 timer, hvilket svarer til ca. 200 km., og herefter er spillover-effekten ikke længere signifikant. Dette bekræfter ifølge Rodríguez-Pose & Crescenzi (2008) antagelsen om, at viden spredes blandt mennesker gennem ansigt-til-ansigt kontakt, hvorfor rejsetiden spiller en rolle, da rejsetiden ikke må være større end hvad en person kan tilbagelægge på en dag (Rodríguez-Pose & Crescenzi, 2008, s. 63).

Den økonometriske model i dette studie operationaliseres gennem variablene F&U-investeringer, som bruges som proxy for spillover af viden, et social indeks der er proxy for

innovationssystemer, som er en regions forudsætninger for at generere ny viden, og BNP per indbygger, der skal give et billede af en regions udviklingsniveau i startåret. Det sociale indeks er en samlet betegnelse for en række variable, som f.eks. arbejdsstyrkens uddannelsesniveau, landbrugets andel af arbejdsstyrken og andelen af unge mellem 15 og 24 år. Disse variable bruges til at beskrive en regions karakteristika, men de bruges også til at beskrive spillover af naboregionernes karakteristika, idet de ganger vækstmatricen på variablene. Ydermere inkluderes naboregioners BNP per indbygger, som en proxy for naboregioners udviklingsniveau. Denne proxy fortæller, om det har en positiv eller negativ effekt på en region, hvis dens naboregioner har en høj velstand. Modellens afhængige variabel er den årlige regionale vækst i BNP per indbygger i perioden 1995-2003, for de europæiske regioner i EU-25, hvor der er data for. Derudover indgår der en række ladedummies, som tager højde for at regioner, der ligger i samme land vil dele samme karakteristika. Inddragelsen af disse dummies er et forsøg på at minimere den rumlige afhængighed i residualerne, der opstår når observationerne er regioner. Modellen estimeres herefter med en heteroskedastisk-konsistent OLS (Rodríguez-Pose & Crescenzi, 2008, s. 55-58).

I modsætning til førnævnte studie går LeSage & Fischers (2008) skridtet videre, og anvender *spatial econometrics* (rumlig økonometri) med henblik på blandt andet at tage højde for rumlig autokorrelation i residualerne. Med rumlig økonometri tages der ikke blot højde for rumlig autokorrelation, men er også en metode hvori spillovers er indbygget i modelspecifikationerne. Der tages her ligeledes udgangspunkt i regionerne i EU-25 og væksten observeret i perioden 1995-2003.

Ifølge forfatterens estimerede model har en region i 's egen udviklingsniveau en positiv effekt på vækst, mens naboregioners udviklingsniveau har en negativ spillover-effekt på region i 's vækst. Den samlede effekt af udviklingsniveauet har en positiv effekt på væksten, men er dog ikke signifikant. De finder ydermere, at regionernes egen patentaktivitet samt naboregioners patentaktivitet har en positiv effekt på vækst. Altså påvirker det en regions vækst positivt, hvis naboregioner producerer ny viden. En regions egen humankapital har en positiv effekt, mens naboregioners niveau for humankapital har en negativ effekt på vækst, og den samlede effekt er ikke statistisk signifikant fra 0. Ifølge LeSage & Fischer (2008) stemmer det også overens med intuitionen om, at det er den intraregionale humankapital der har den største betydning for væksten, mens niveauet i humankapitalen i naboregioner ikke vil have en særlig stor effekt, hvis

nogen, på regionens egen vækst. De finder desuden at befolkningstæthed har en stor negativ total effekt på væksten, da effekten af både regioners og deres naboers befolkningstæthed er negativ. De finder også, at en regions industrisammensætning har en stor betydning for væksten. De finder at industrier, som bl.a. landbrug-, fødevare- og byggebranchen, har en negativ effekt på væksten, mens industrier som inkluderer detailhandel, hotel og restaurationsbranchen og transport har en positiv effekt på væksten. Dette indikerer at en regions vækstpotentiale afhænger af, hvilke industrier der er dominerende i regionen.

Resultaterne er udledt fra en modelspecifikation kendt som en Spatial Durbin Model, som både modellerer effekterne af regionernes egen produktion af viden og spillover-effekten af naboregionernes produktion af viden, samt andre regioners vækst.⁶ Resultaterne er udledt på baggrund af tre forskellige vægtmatricer. Disse er baseret på tre former for distance, herunder distance i direkte fugleflugtslinje, distance i rejsetid på vej og rejsetid for godstransport.

Den afhængige variabel er den gennemsnitlige bruttoværditilvækst per indbygger i perioden 1995-2003. Derudover består modellen af 23 uafhængige variable og deres laggede værdier, der alle er målt i periodens startår. Disse variable omfatter blandt andet; udviklingsniveauet, der er målt bruttoværditilvæksten per indbygger; humankapital, som er målt i uddannelsesniveaue for befolkning i alderen 15 og over; patentaktiviteter, der er målt i antal totale patentansøgninger; befolkningstæthed, målt ved befolkningstæthed per km²; og regionens industrisammensætning, målt ved beskæftigelse indenfor en række industrier.

I et studie af Basile, Capello & Caragliu (2011) testes ved hjælp af rumlig økonometri, hypotesen om at geografisk nærhed ikke nødvendigvis er den eneste måde at fange mekanismerne bag spillover af viden. Andre alternativer kunne være institutionel nærhed eller teknologisk nærhed. I dette studie tester de den fælles effekt af geografisk og relationel nærhed, hvor relationel nærhed måles ved forskellen i tilliden mellem to regioner, ud fra den baggrund at økonomiske agenter agerer forskelligt, og at den har en indflydelse på læringsprocessen. Hypotesen testes bl.a. ved en Spatial Durbin model, og stikprøven består af 249 NUTS 2 regioner i EU-27 i perioden 1990-2004 (Basile, Capello, & Caragliu, 2011, s. 22).

Studiet konkluderer at både geografisk og relationel nærhed er afgørende for spillover af viden og at interaktionen mellem disse to nærhedsprincipper også er signifikant. De finder at spillover af viden falder jo længere væk to regioner er fra hinanden, hvilket ifølge Basile et al. (2011)

⁶ Denne modelspecifikation anvendes også i dette projekt, og er nærmere forklaret i afsnit 4.3.1.

stemmer overens med andre lignende studier, og at den geografiske distances rolle ikke bliver nævneværdig ændret ved at kontrollere for relationel distance. Interaktionen viser at regioner der er naboer i geografiske termer har en større vidensudveksling, hvis de samtidig også ligger tæt på hinanden i relationelle termer. Dette betyder at bare fordi to regioner er geografiske naboer, så sker der ikke nødvendigvis en automatisk spillover af viden, da det også har stor betydning, hvordan forholdet mellem de to regioner er.

Studiets afhængige variabel er produktivitetsvæksten i perioden 1990-2004, som er målt i forholdet mellem den reale værditilvækst og beskæftigelse. De uafhængige variable består af opsparingsraten, der måles som bruttoinvesteringers gennemsnitlige andel af bruttoværditilvæksten; akkumulationen af humankapital, målt i hvor stor en andel af arbejdsstyrken der har en sekundær uddannelse; væksten i beskæftigelsen; samt naboregionernes værdier af førnævnte variable, som et udtryk for spillovers. Der arbejdes med en distancematrice med en øvre grænse på 424 km., da det er ved denne distance at alle regioner i EU-27 har en nabo (Basile, Capello, & Caragliu, 2011, s. 26).

3.2 Eget bidrag

Der er mange studier, som beskæftiger sig med regional vækst, og disse modellerer vækst på hver sin måde, men de overlapper i mange henseender. Dette speciale bygger på tidligere studiers indsigter og relateret litteratur, hvorfor specialet i en høj grad også vil overlappe med tidligere studier. Dette speciale adskiller sig dog fra andre studier, idet videns tavse og kodificeret komponent specificeres separat i den økonometriske model. Mange studier refererer til disse to komponenter af viden, heriblandt Rodríguez-Pose & Crescenzi (2008), men de bruger blot videns tavse komponent til at argumentere for, at spillovers af viden ikke rækker langt, og anvender derefter F&U-investeringer, som en paraply-proxy for spillovers af viden. I nærværende tekst anvendes i stedet to proxies, som forventes at opfange hhv. den kodificerede og tavse komponent af viden. Disse proxies er hhv. antal patentansøgninger og antal ansatte indenfor Science & Technology-sektoren, og argumenterne herfor præsenteres i afsnit 4.5.1 Variabelbeskrivelse. Den geografiske dimension modelleres i Rodríguez-Pose & Crescenzi (2008) som sagt vha. en invers kontinuer distancematrice, således at fjerne regioner vægtes mindre end nære regioner. I dette speciale modelleres den geografiske dimension i stedet vha. to vægtmatricer, som hver specificerer et distancebånd, og disse bånd er valgt med udgangspunkt i hvad tidligere studier finder, er rækkevidden for spillover-effekter.

4 Metode

Når et datasæt består af observationer fra f.eks. regioner inden for et geografisk område kan der opstå *spatial dependence*, dvs. rumlig afhængighed, fordi der vil være en tendens til at værdier for en region vil have ligheder, pga. gensidig afhængighed, med de regioner der ligger tæt på og især regioner indenfor samme land (LeSage & Pace, Introduction to Spatial Econometrics, 2009, s. 2). Med rumlig økonometri tages der dog højde for netop dette problem, ved at inddrage denne afhængighed, og derved undersøge hvor meget andre regioner tæt på kan påvirke egen performance. Derfor er rumlig økonometri særlig velegnet til undersøgelser, hvor eventuelle spillover-effekter mellem regioner er genstandsfelt. Rumlig økonometri udgør et delfelt af økonometriske metoder, som er optaget af rumlige aspekter som er til stede i tværsnits- og rumtids observationer (Anselin, 2010, s. 5). De første artikler i feltet blev udgivet i 1970'erne, og forfattet af regionale udviklingsforskere, artiklerne var dog fraværende i de store økonomiske og økonometriske tidsskrifter (Anselin, 2010, s. 7). Situation har sidenhen ændret sig, og rumlig økonometri har fundet vejen til mainstream tidsskrifter, og er endda blevet adopteret i empiriske studier af vækst, men feltet har endnu ikke vundet fuldt indpas i de mest populære økonometriske lærebøger (Anselin, 2010, s. 8).

4.1 Datastruktur og den ikke-rumlige model

Studier af regional vækst anvender typisk enten tværsnitsdata eller paneldata. Tværsnitsdata er kendetegnet ved at indeholde observationer af flere enheder (lande, regioner mm.) over en tidsperiode, mens paneldata indeholder observationer over flere enheder over flere tidsperioder. I nærværende studie benyttes tværsnitsdata over regionerne i EU-15. Dette er på baggrund af ønsket om at undersøge regionale vækstrater over en periode. Alternativt kunne der ses på flere tidsperioder ved at benytte paneldata, men pga. datamangel giver det et ubalanceret paneldatasæt. Data mangler af flere grunde. En årsag er at regionerne er begyndt at registrere data på forskellige tidspunkter, og en anden årsag er at regionsgrænserne har ændret sig, fordi nogle administrative enheder bliver lagt sammen eller opdelt i mindre enheder, hvilket giver databrud. Eksempelvis bestod London af to NUTS2 regioner før 2010, men herefter blev den yderligere opdelt i fem NUTS2 regioner. Ulempen ved at anvende et ubalanceret datasæt er, at usikkerheden om analysens resultater vil stige. Derudover tager paneldata også højde for tidsdimensionen, men da der ikke er tilstrækkeligt med data over så lang en årrække, vil tværsnitsdimensionen stadig være dominerende, hvorfor det i nærværende projekt vil være overflødig at beskæftige sig med paneldata, når tidsdimensionen alligevel ikke vil indgå.

Indledningsvist opstilles en multipel lineær regressionsmodel af formen:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u \quad (6)$$

Denne estimeres ved hjælp af ordinary least squares (OLS) metoden. Denne metode går ud på at finde den løsning der er BLUE, hvilket betyder at estimerne både opfylder Gauss-Markov antagelserne og samtidig har mindst varians. Gauss-Markov antagelserne er fem antagelser en model skal overholde for at kunne betegnes som BLUE (Wooldridge, 2013, s. 71-74, 81). 1) modellen skal kunne skrives som vist i ligning 6, hvor betaværdierne er de ukendte parametre, som forsøges estimeret, mens u er et uobserveret fejld. 2) Stikprøven skal være tilfældig og skal bestå af n observationer, der følger modellen i foregående antagelse. 3) Der må ikke være perfekt kollinearitet, dvs. at modellens uafhængige variable ikke må have en perfekt lineær sammenhæng. 4) Fejlledets betingede middelværdi skal være 0, og den forventede værdi af fejleddet er 0 for alle værdier af de uafhængige variable. 5) Fejlledene skal være homoskedastiske, dvs. at de skal have samme varians for alle værdier af de uafhængige variable. Derudover må fejleddet ikke være afhængig af de uafhængige variable, da det vil resultere i heteroskedastiske fejld, som gør fejleddet biased. Overholdes antagelserne ikke, skal alternative metoder tages i brug alt efter hvilke antagelser der brydes. Da der i dette projekt arbejdes med regionale data, vil det være forventeligt at der er autokorrelation mellem residualerne for regioner i samme land, da regioner i samme land deler visse karakteristika, såsom sprog, kultur samt samme love og regler. Er dette tilfældet, vil det være et udtryk for geografisk afhængighed som bryder med Gauss-Markov antagelse 3). Dette problem kan behandles ved hjælp af metoder anvendt i *rumlig økonometri*.

4.2 Rumlig afhængighed

Det første skridt i brugen af rumlig økonometri er, at tjekke om der overhovedet er rumlig afhængighed mellem residualerne i den lineære ikke-rumlige model. For at teste for rumlig afhængighed udføres en Moran's I test, som vil forklares senere i afsnittet. Indledningsvist kræver denne testmetode, at der konstrueres en vægtmatrice.⁷ Vægtmatricen kan konstrueres på adskillige måder, og alt efter genstandsfelt og datatilgængelighed må den der passer bedst vælges. I nærværende projekt konstrueres der en distancematrice ud fra punktkoordinater (x,y) for de regionale geometriske tyngdepunkter i et kartesisk todimensionelt koordinatsystem. Et geometrisk tyngdepunkt er et udtryk for regionens geografiske centrum.

⁷ Vægtmatricen skal bruges til at vægte regionernes påvirkning på hinanden.

Distancen beregnes mellem to geometriske tyngdepunkter ud fra den Euclidean distance ligning:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (7)$$

Dette resulterer i en $n \times n$ matrice der indeholder distancen mellem region i og j . Denne måde at definere en distancematrice på er også brugt i andre studier, se f.eks. (Charlot, Crescenzi, & Musolesi, 2014). For bedre at kunne forholde sig til distancen skal det omregnes til km. Der er ikke nogen specifik enhed for den Euclidean distance, fordi det er afstanden mellem to koordinater i et koordinatsystem. Der er ikke nogen direkte måde at omregne den Euclidean distance om til km., derfor bruges Haversine-formlen (Robusto, 1957). Med Haversine-formlen i ligning 8 kan der ud fra bredde- og længdegradskoordinater beregnes afstanden mellem to punkter:

$$d = 2r \arcsin \left(\sqrt{\sin^2 \left(\frac{\varphi_2 - \varphi_1}{2} \right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right) \quad (8)$$

Hvor φ er breddegraden for region i og j og λ er længdegraden for region i og j . R er jordens radius, her opgjort i km. Da jorden ikke er helt rund vil dens radius være forskellig alt efter, hvor på jorden regionerne befinder sig, hvorfor der bruges en middelværdi som svarer til 6.371 km, hvilket også betyder at der er en lille afvigelse i forhold til den egentlige distance. Når afstanden er beregnet kan forholdet mellem Haversine-distancen og den Euclidean distance findes, og derved give et ca. mål for hvor mange km. den Euclidean distance udgør. I det følgende vil der komme et eksempel på, hvordan dette gøres.

Tages der udgangspunkt i region $AT11$ og $AT12$ er den Euclidean distance 1,1069385. Bredde- og længdegradskoordinaterne for region $AT11$ og $AT12$ er hhv. (47,5022;16,4368) og (48,2454;15,7607) (Eurostat, 2015). Indsættes koordinaterne i ligning 8, fås $d = 7554,055$ m, hvilket svarer til 75,54 km. Herefter findes forholdet mellem Haversine-distancen og den Euclidean distance, og resultatet er at en enhed af den Euclidean distance svarer til ca. 68,24 km. Denne beregnede distance bruges nu til at definere indenfor, hvilken distancegrænse en region betegnes som en naboregion. I nærværende projekt konstrueres to distancematrixer til at undersøge, hvornår spillover-effekten ikke længere er til stede. Den første distancematrice, W_1 , har en distancegrænse på $= 3$, hvilket svarer til ca. 205 km. jf. ovenstående. Denne grænse er valgt, fordi tidligere studier har vist at for europæiske regioner falder effekten af naboregioner

efter ca. 200 km. (Rodríguez-Pose & Crescenzi, 2008, s. 63). Den anden distancematrice, W_2 , konstrueres på samme måde, men med en distance på $3 > d < 5$, hvor 5 svarer til ca. 340 km.

Ud fra de valgte distancegrænser konstrueres en binær matrice, w_{ij} , og hvis distancen mellem region i og j er indenfor distancegrænsen tildeles den værdien 1, og er den ikke tildeles den værdien 0 (LeSage & Pace, 2009, s. 9).

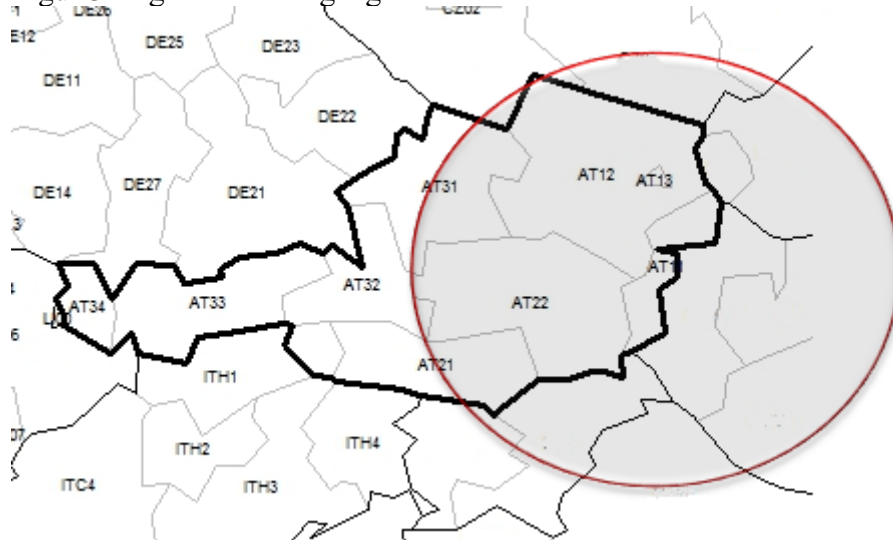
$$w_{ij} = \begin{cases} 1 & , \quad 0 \leq d_{ij} \leq d \\ 0 & , \quad d_{ij} > d \end{cases} \quad (9)$$

Dette giver følgende binær matrice:

	AT11	AT12	...	UK11
AT11	0	1	...	0
AT12	1	0	...	0
⋮	⋮	⋮	⋮	⋮
UK11	0	0	...	0

I elementet hvor AT11 og AT12 krydser indikerer 1-tallet, at de to regioner er naboer, mens 0 i elementet hvor AT11 og UK11 krydser indikerer at de to regioner ikke er naboer. Da en region ikke kan være nabo til sig selv, vil elementerne i diagonalen bestå af 0'er. I Figur 3 er der et kort over Østrig, og figuren illustrerer grafisk hvor mange naboregioner den østrigske region AT11 (Burgenland) får, når distancegrænsen er ca. lig 205 km. Det fremgår af figuren, at AT11 per definition får fem naboregioner.

Figur 3 Region AT11's geografiske naboer illustreret



Note: Dette er ikke en eksakt afbildning af en afstand på 205 km. i radius.
 Kilde: Egen tilvirkning, baseret på Eurostat (2015).
 © EuroGeographics for de administrative grænser.

En af begrænsningerne ved denne metode er, at distancen er mellem geometriske tyngdepunkter, dvs. for at en region kan defineres som en nabo kræver det, at den anden regions geometriske tyngdepunkt ligger indenfor distancegrænsen. Dette betyder at en region i praksis godt kan have fysiske grænser med en anden region, men fordi dens geometriske tyngdepunkt ikke ligger indenfor distancegrænsen, vil den ikke tælle med som en nabo. Derfor vil nogle regioner have flere geografiske naboer end, hvad denne metode opfanger. Dette betyder at spillover-effekten fra de regioner der ikke indgår som naboer ikke opfanges, selvom de også må antages at have en effekt. Det egentlige problem opstår, hvis en region, pga. distancegrænsen, ikke kommer til at tælle med, selvom det er den naboregion, hvor den største spillover-effekt kommer fra. Et alternativ kunne være at definere naboer efter om regionerne deler grænser. Begrænsning ved dette alternativ er, at spillover-effekten ikke nødvendigvis kun sker til grænsenaboer, men kan også ske til 2.-ordens naboer, som er naboens nabo.

Næste skridt er at rækkestandardisere matricen. Dette sker ved at hvert 1-tal i rækken divideres med det samlede antal 1-taller i rækken, og dette skulle gerne resultere i at hver række summere til 1. Således fås den endelige vægtmatrice, der ser ud som følger:

	<i>AT11</i>	<i>AT12</i>	...	<i>UK11</i>
<i>AT11</i>	0	$\frac{1}{5}$...	0
<i>AT12</i>	$\frac{1}{6}$	0	...	0
⋮	⋮	⋮	⋮	⋮
<i>UK11</i>	0	0	...	0

Da *AT11* i alt har 5 naboregioner, vil alle regioner have vægten $1/5$, intuitivt betyder det at hver region vil tilskrives $1/5$ af den samlede interregionale påvirkning på denne østrigske region. Regioner som ikke er naboer vil stadig være 0. Når den standardiserede vægtmatrice konstrueres på denne måde, vil alle regioners naboer have samme vægt. Dette resulterer i en vektor, der giver den gennemsnitlige effekt naboregioner har på en region (LeSage & Pace, 2009, s. 10). Implikationen af gennemsnitseffekten er, at spillover-effekten fra stærke regioner bliver udlignet af effekten fra mindre stærke regioner. Begrænsningen ved denne konstruktion er at det ikke bliver muligt at fange at nogle naboregioner har en større spillover-effekt end andre. Denne metode tillader desuden ikke regioner, der ingen naboer har. Det betyder at de regioner der ud fra specifikationen ikke har nogen nabo, må tages ud. Dette problem opstod i to tilfælde

herunder ved øen Sardinien⁸, hvor afstanden til fastlandet var større end distancegrænsen, samt ved den nordligste region i Sverige, som er så stor at distancegrænsen ikke når ud af regionsgrænsen, hvorfor disse regioner bliver udeladt.

Når vægtmatricen er konstrueret kan det nu tjekkes om der er rumlig autokorrelation mellem residualerne. Dette gøres som sagt ved hjælp af en Moran's I test, som er defineret for en residual vektor x :

$$I = \frac{n}{S_0} \frac{x'Wx}{x'x} \quad (10)$$

Her er x en $n \times 1$ vektor af uafhængige variable, W er en $n \times n$ vægtmatrice, n er antallet observationer, i dette tilfælde antallet af regioner og S_0 er summen af elementerne fra W (Liu, Griffin, & Kirkpatrick, 2014, s. 14). Ud fra denne ligning kan nulhypotesen om, at der ingen rummelig autokorrelation er, testes. Moran's I skal fortolkes som en korrelationskoefficient. Positive Moran's I værdier indikerer at stærke regioner har en tendens til at koncentrere sig sammen, mens svagere regioner koncentrerer sig sammen. En negativ Moran's I værdi indikerer at stærke regioner er omgivet af svagere regioner, og derved ikke har stærke naboer omkring sig (Liu, Griffin, & Kirkpatrick, 2014, s. 14). Hvis testen viser at der er rumlig autokorrelation, skal alternative regressionsmetoder til OLS bruges.

4.3 Rumlig økonometri

Der er ikke dedikeret meget plads til rumlig økonometri i traditionelle økonometriske lærebøger. LeSage & Pace (2009) har til gengæld udgivet en bog, som alene er fokuseret på rumlig økonometri og denne har en udbredt anvendelse, og det er denne der hovedsageligt trækkes på i nærværende afsnit. Metodologien indenfor dette felt omfatter typisk fire elementer, herunder modelspecifikation, -estimation, specifikationstest samt forecast (Anselin, 2010, s. 6). Med hovedfokus på de to første elementer præsenteres indledningsvist valg af modelspecifikation og herunder motivationen bag, efterfulgt af en begrundelse for valget af estimationsmetode, og en præsentation af hvordan estimerne tolkes.⁹

4.3.1 Modelspecifikationer

Modelspecifikation omhandler det formelle matematiske udtryk for den rumlige afhængighed i regressionsmodellen (Anselin, 2010, s. 6). I nærværende studie tages der udgangspunkt i en SDM-model og en SDEM-model.

⁸ Dette skyldes dog kun at Korsika er udelukket pga. datamangel, hvilket ikke efterlader Sardinien nogen naboer.

⁹ Forecast ses der bort fra.

Spatial Durbin modellen

Udgangspunktet for dette studie er modelspecifikationen kendt under navnet *Spatial Durbin Model* (SDM), som antager formen udtrykt i ligning (11):

$$y = \rho W_y + \alpha \iota_n + X\beta + WX\gamma + \varepsilon \quad (11)$$

Den afhængige variabel y repræsenterer en $n \times 1$ -vektor, mens X er en $n \times k$ matrix, der indeholder n observationer (regioner) og k uafhængige variable, mens ι repræsenterer en $n \times 1$ vektor af 1-taller associeret med skæringspunktet, α (LeSage & Fischer, 2008, s. 278; LeSage & Pace, Introduction to Spatial Econometrics, 2009, s. 46). Størrelsen γ refererer til størrelsen på spillover-effekten af X på y . Skalarparameteren, ρ , måler hvor stor en effekt andre regioners y har på region i 's y (LeSage & Fischer, 2008, s. 278).

Således repræsenterer y i nærværende tekst en vektor, der indeholder n observerede størrelser for væksten i BNP per indbygger i de respektive regioner i stikprøven. Ydermere dækker X over de uafhængige variable, som antages at påvirke væksten. Forskellen fra den almindelige lineære regressionsmodel jf. ligning (6) er, at den rumlige dimension inkorporeres i SDM vha. W , som er en $n \times n$ ikke-stokastisk og ikke-negativ vægtmatrix, som blev konstrueret i forrige afsnit 4.2. Elementerne i W anvendes til at specificere den rumlige afhængighedsstruktur blandt observationerne (LeSage & Fischer, 2008, s. 278). Såfremt observation for region i er relateret til region j , så vil $W_{ij} > 0$, ellers vil $W_{ij} = 0$. Det særlige ved SDM-modellen i relation til andre rumlige modeller er, at den inkluderer både en vektor med et rumligt lag af den afhængige variabel, Wy , og et rumligt lag af de uafhængige variable, $WX\gamma$. Det betyder i praksis, at der både tages højde for effekten af alle regioners afhængige variabel, y , og naboregioners uafhængige variable, X , på region i 's afhængige variabel. Teknisk betyder det at Wy indeholder en lineær kombination af vækstraterne i alle regionerne (defineret ud fra hvor $W_{ij} > 0$), og vektoren fanger således en rumlig afhængighed i den afhængige variabel (LeSage & Fischer, 2008, s. 278). ρ er som sagt en skalarparameter, og måler i denne kontekst effekten af alle regioners vækstrater på region i 's vækstrate, og antages at være positiv og have en værdi mindre end 1, således at den lineære kombination af alle andre regioners vækstrater har en positiv effekt på region i 's vækstrate (LeSage & Fischer, 2008, s. 278). Hvis ρ er lig 0, så er implikationen, at andre regioners vækstrater ikke har nogen effekt på region i 's vækst. Matrix WX indeholder en lineær kombination af forklarende variable fra naboregioner. Størrelsen ε svarer til en

fejlledsvektor med en dimension på $n \times 1$, og $\varepsilon \sim N(0, \sigma_\varepsilon^2)$, dvs. den følger en normalfordeling og har en konstant varians (LeSage & Fischer, 2008, s. 279).

Spatial Durbin Error modellen

Spatial Durbin Error modellen (SDEM) inkluderer ligesom SDM et rumligt lag af de forklarende variable i form af WX , men til forskel fra sidstnævnte inkluderer SDEM ikke et separat rumligt lag af de afhængige variable (LeSage & Pace, 2009, s. 41-42). SDEM har derimod åbnet op for rumlig afhængighed i fejlleddene. En SDEM kan således ifølge LeSage & Pace (2009, s. 42) udtrykkes som i ligning (12):

$$y = X\beta + WX\gamma + \iota_n\alpha + u \quad (12)$$

$$u = R^{-1}\varepsilon = (I_n - \lambda W)^{-1}\varepsilon \quad (13)$$

Her er y igen den afhængige variabel repræsenteret ved en $n \times 1$ -vektor, mens X er en $n \times k$ matrix, der indeholder n observationer (regioner) og k uafhængige variable, mens ι endnu engang repræsenterer en $n \times 1$ vektor af 1-taller associeret med skæringspunktet, α . Størrelsen u refererer til fejleddet, og er udtrykt i ligning (13). I ligning (13) er λ en skalarparameter, som måler hvilken effekt en ændring i en regions fejllid har på andre regioners fejllid (LeSage, 2014, s. 16).

Motivation

Motivationen til at benytte SDM og SDEM beror for det første på, at denne specifikation tillader at der kan inkorporeres teoretiske aspekter ind i modellen, som ikke er mulige ved en almindelig lineær model. Valget af SDM- og SDEM-modellen sker således af hensyn til at kunne modellere spillovers af de forskellige typer af viden (tavs og eksplicit) i modellen for at kunne fange deres potentielle effekt på vækst ved hjælp af $WX\gamma$. Ydermere muliggør SDM, at der samtidigt kan kontrolleres for effekten af andre regioners vækstrater ved hjælp af ρWy , mens SDEM kontrollerer for andre regioners fejllid. Relativt til andre rumlige modelspecifikationer, er det kun SDM der både inkluderer et rumligt lag af den afhængige variabel og af de uafhængige variable, mens SDEM er den eneste der inkluderer både et rumligt lag af de uafhængige variable og fejleddet.¹⁰ Dermed vil den regionale vækst i perioden 2002-2014 modelleres på to måder i

¹⁰ Der er andre rumlige modelspecifikationer, som blot inkorporerer visse aspekter af SDM (og SDEM). Der eksisterer en *spatial autoregressive model* (SAR), som inkluderer et rumligt lag af den afhængige variabel, y , i form af ρWy . En anden model, kendt som *spatial lag of X* (SLX) inkluderer derimod blot et rumligt lag af de afhængige variable, X , i form af WX . Ydermere findes en *spatial error model* (SEM), hvor der blot er inkluderet et

dette projekt. Den vil i begge tilfælde være afhængig af regioners egne karakteristika og spillover-effekter i de forklarende variable, men de vil adskille sig ved at den ene tager højde for andre regioners vækstrater, mens den anden tager højde for andre regioners fejlede. Der er et teoretisk belæg for at operere med en såkaldt vækstspillover, da Capello (2009, s. 640) argumenterer for, at en regions vækst kan påvirkes af naboregioners vækst gennem handels- og markedsforbindelser.

Valget af SDM og SDEM-modellen bygger for det andet på økonometriske hensyn. Der eksisterer altid en risiko for, at der er udeladte forklarende variable, som korrelerer med en af de inkluderede forklarende variable i modellen, hvilket kan skabe bias i parameterestimerne, dvs. visse af parametrene kan blive over- eller underestimeret. Udeladte variable giver dog større bias i OLS-estimer, end de gør i SDM-modellers estimer, hvis der samtidig er en rumlig afhængighed i de uafhængige variable og de afhængige variable (LeSage & Pace, 2009, s. 67). Hos OLS-estimer vil størrelsen af bias være afhængig af i hvilken grad der er rumlig afhængighed mellem de inkluderede uafhængige variable eller fejleddet. Hos SDM vil der kun være bias i β , og denne bias vil være konstant og kun være afhængig af styrken af relationen mellem den inkluderede forklarende variabel og den udeladte variabel (LeSage & Pace, 2009, s. 67). Ved at tage udgangspunkt i en SDM, så vil styrken af de udeladte variables bias ikke være afhængig af den rumlige afhængighed, og derfor vil estimerne for en rumlig økonometrisk model være mindre påvirket af udeladte variable i forhold til OLS-estimer, og i dette henseende er SDM mere favorabel (LeSage & Pace, 2009, s. 67-68). I nærværende studie er fysisk kapital ikke medtaget som en forklarende variabel, pga. utilstrækkelig data, og udgør derfor i Romers (1990) objektiv en udeladt variabel. Ved at vælge SDM, afbødes der for alt for store bias i parameterestimerne, fordi denne model tager højde for rumlig afhængighed. SDM-modellen anbefales desuden stærkt af LeSage & Pace (2009, s. 156), som mener at den ikke er diskuteret nok i økonometrilitteraturen, når det handler om valget af modelspecifikation. Argumentet for SDM gør sig også gældende for SDEM, men her er det i stedet modelleringen af rumligheden i fejleddene, der kan afbøde for store bias ved udeladte variable (Cook, Hays, & Franzese, 2015, s. 16).

rumligt lag af fejleddet (LeSage & Pace, Introduction to Spatial Econometrics, 2009, s. 16, 30-32). For at se flere andre rumlige modelspecifikationer, se (LeSage & Pace, Introduction to Spatial Econometrics, 2009).

4.3.2 Modelestimation og fortolkning

Når den rumlige afhængighed er inkorporeret i regressionsspecifikationen, så må en passende estimationsmetode anvendes til, at tage højde for konsekvenserne ved den rumlige afhængighed (Anselin, 2010, s. 6). Anvendelse af Ordinary Least Squares-estimationsmetoden har vist sig at lede til inkonsistente estimater for modellens regressionsparametre, når rumlig afhængighed er til stede, og især hvis der er udelade variable som korrelerer med modellens forklarende variable (LeSage & Pace, 2009, s. 45). Maximum Likelihood estimation (MLE) har derimod vist sig at give konsistente resultater for rumlige modelspecifikationer, heriblandt SDM-modellen (LeSage & Pace, 2009, s. 45), hvorfor det er denne er benyttet.

Fortolkning af koefficienter (parameterestimer)

Det er ikke muligt at lave en ligefrem fortolkning af en SDM-models parameterestimer udledt vha. MLE, fordi de bygger på et rigere sæt af informationer om forholdet mellem observationerne, da der simultant tages højde for informationer om regioner og deres naboer. Den lineære models estimer har derimod en ligefrem fortolkning, og tolkes for region i som den partielt afledte effekt på den afhængige variabel y_i , skabt pga. en ændring i den uafhængige variabel x_{ir} :

$$\frac{\partial y_i}{\partial x_{ir}} = \beta_r \text{ for alle } i, r \quad (14)$$

Den lineære model indeholder ikke informationer om region i 's naboer, hvorfor $\frac{\partial y_i}{\partial x_{jr}} = 0$ for $j \neq i$ og alle uafhængige variable, r (LeSage & Pace, 2009, s. 34). Grundet SDM-modellens rumlige lag af den afhængige variabel og de uafhængige variable, så vil en ændring i enhver given uafhængig variabel for region i have en *direkte effekt* på region i selv, såvel som en *indirekte effekt* (spillover-effekt) på andre regioner $j \neq i$ (LeSage & Pace, 2009, s. 33). Dette betyder, at udtrykket for den lineære models estimer ikke holder for SDM-modellen. Der findes udtryk for både den direkte og indirekte effekt, der i essensen ligner ovenstående, og disse kan udledes med udgangspunkt i en omskrivning af SDM-modellen udtrykt i ligning (15), hvilket giver udtrykket i ligning (16):

$$(I_n - \rho W)y = X\beta + WX\theta + I_n\alpha + \varepsilon \quad (15)$$

$$y = \sum_{r=1}^k S_r(W) x_r + V(W)I_n\alpha + V(W)\varepsilon \quad (16)$$

Modellen indeholder stadigvæk k uafhængige variable, hvor x_r angiver den r 'te uafhængige variabel¹¹, og $S_r(W)$ er en multiplikatormatrix, som er udtrykt ved:

$$S_r(W) = V(W)(I_n\beta_r + W\theta_r)$$

$$V(W) = (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots$$

Hvor I_n svarer til en $n \times n$ identitetsmatrix, β svarer til en $k \times 1$ vektor, der indeholder regressionsparametrene associeret med de forklarende variable, mens θ svarer til en $k \times 1$ vektor, der indeholder regressionsparametrene associeret med de rumligt laggede uafhængige variable, WX (LeSage & Fischer, 2008, s. 283). I matrixen $S_r(W)$ er det muligt at spore de direkte og indirekte effekter, og dette kan lettest illustreres ved indledningsvist at ekspandere (16), hvilket giver ligning (17) præsenteret foruden (LeSage & Pace, 2009, s. 35):

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{r=1}^k \begin{pmatrix} S_r(W)_{11} & S_r(W)_{12} & \dots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & & \\ \vdots & \vdots & \ddots & \\ S_r(W)_{n1} & S_r(W)_{n2} & \dots & S_r(W)_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ x_{3r} \\ x_{4r} \end{pmatrix} + V(W)l_n\alpha + V(W)\varepsilon \quad (17)$$

Et givet element i $S_r(W)$ -matrixen defineres efter hhv. hvilken række og hvilken søjle den befinder sig i, dvs. det i, j 'te element af matrixen $S_r(W)$ noteres $S_r(W)_{ij}$. Tages der udgangspunkt i en enkelt observation for y , så bliver det mere gennemskueligt, hvorfor de direkte og indirekte effekter fremgår heraf:

$$y_i = \sum_{r=1}^k [S_r(W)_{i1}x_{1r} + S_r(W)_{i2}x_{2r} + \dots + S_r(W)_{in}x_{nr}] + V(W)_{i}l_n\alpha + V(W)_{i}\varepsilon$$

Det følger af ovenstående ligning, at den partielt afledte effekt på y_i mht. x_{jr} potentielt er *ikke-nul* i en SDM-model, hvilket per definition ikke er mulig i den lineære model når $j \neq i$, fordi der ikke haves informationer om naboregioner (LeSage & Pace, 2009, s. 35). Den partielt afledte effekt antager i SDM en værdi, som er bestemt af det i, j 'te element af $S_r(W)$ -matrixen.

Implikationen er således, at en ændring i den r 'te variabel i en SDM-model har en partielt afledt effekt på y_i , som er lig det i, j 'te element af en $n \times n$ matrixe, $S_r(W)$:

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij} \quad (18)$$

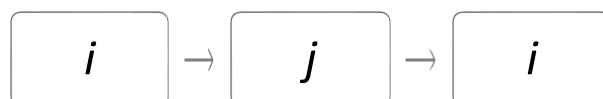
¹¹ Hvilket svarer til den r 'te søjle i X .

Ligning (18) er således et udtryk for den *indirekte effekt / spillover effekt*, idet ændringen i regions i 's afhængige variabel kommer fra en ændring i region j 's forklarende variabel (LeSage & Pace, 2009, s. 35). En ændring i en uafhængig variabel for en enkelt region kan således potentielt påvirke den afhængige variabel hos alle andre regioner. Den direkte effekt kan ligeledes findes i $S_r(W)$ -matricen, i elementet $S_r(W)_{ii}$, som måler den partielle effekt på den i 'te regions afhængige variabel, y_i , forårsaget af en ændring i den i 'te regions egen uafhængige variabel, x_{ir} :

$$\frac{\partial y_i}{\partial x_{ir}} = S_r(W)_{ii} \quad (19)$$

Den direkte effekt omfatter også *feedback loops*, hvor observation i påvirker observation j , som også påvirker observation i , som illustreret i Figur 4.

Figur 4 Feedback loops



Kilde: Egen tilvirkning, baseret på (LeSage & Pace, 2009, s. 35-36).

Denne feedback skyldes, at effekterne passerer igennem naboregioner og når tilbage til regionen selv, idet en region også udgør sin nabos nabo, hvorfor en ændring hos region i vil påvirke region j , som vil skabe højere ordens feedback effekter på region i selv (LeSage & Pace, 2009, s. 36). En region i kan således betragtes som sin egen andenordens nabo. Styrken af feedback-effekten afhænger af 1) regionernes placering i rummet, 2) graden hvormed regionerne er forbundet, som indikeret af modellens vægtmatrice, W , 3) samt parameteren ρ , som måler styrken af den rumlige afhængighed, og 3) ikke mindst af parameterne β og θ (LeSage & Pace, 2009, s. 36).

De diagonale elementer i $n \times n$ matricen, $S_r(W)$, repræsenterer de direkte effekter, mens de off-diagonale elementer repræsenterer de indirekte elementer (LeSage & Pace, 2009, s. 36). Den direkte plus den indirekte effekt er, alt andet lige, lig den totale effekt ved en ændring i den r 'te variabel.

Effekten af ændringen i en uafhængig variabel vil variere fra region til region, hvorfor Pace & LeSage (2006) foreslår, at der anvendes tre skalærer som deskriptive mål for regionernes varierende effekter (LeSage & Pace, 2009, s. 36).

- *Den gennemsnitlige direkte effekt* konstrueres ved hjælp af et gennemsnit af de diagonale elementer i $S_r(W)$, dvs. der findes et gennemsnit af de direkte effekter, $S_r(W)_{ii}$. Dette mål beskriver således den gennemsnitlige effekt på y_i , forårsaget af en ændring i den r 'te variabel, $x_{i,r}$, for alle i observationer.¹²
- *Den gennemsnitlige indirekte effekt* konstrueres ved at finde gennemsnittet af off-diagonale elementer i $S_r(W)$. Først findes der et gennemsnit af rækkeelementerne, og dernæst findes der et gennemsnit af rækkernes gennemsnit.
- *Den totale gennemsnitlige effekt* svarer blot til summen af den direkte og indirekte effekt. Det betyder i praksis, at der findes summen af diagonale og off-diagonale elementer i $S_r(W)$, hvorefter der findes et gennemsnit af disse.

Disse tre deskriptive mål vil derfor beregnes med henblik på at kunne anvende dem til at knytte en kommentar til SDM-modellens uafhængige variable, hvad angår størrelsesorden, fortegn og signifikans, efter modellen er estimeret. For at kunne konkludere noget om den statistiske signifikans af den direkte, indirekte (spillovers) og totale effekt, så er det nødvendigt videre at have fordelingen for hver effekts skalarmål. LeSage & Pace (2009, s. 39) foreslår, at der simuleres en fordeling af direkte og indirekte effekter, hvorudfra der kan estimeres en spredning for effekterne. Hvis estimererne i SDM-modellen er signifikante, så betyder det således ikke, at de automatisk vil have en signifikant direkte og indirekte effekt.

En SDEM har derimod en mere ligefrem fortolkning, idet de direkte effekter blot er repræsenteret ved modellens parameter β , mens de indirekte effekter svarer til γ (LeSage & Pace, 2009, s. 42). Disse ligefremme fortolkninger gør det muligt at anvende spredningsmål, såsom standardafvigelse og t-statistikker som basis for at udlede signifikans af den direkte og indirekte effekt, hvorfor signifikansniveauerne ikke behøver at blive udledt fra en simulering, ligesom hos SDM (LeSage & Pace, 2009, s. 42).

¹² Bemærk, at dette deskriptive mål essentielt ligner den lineære models mål for regressionskoefficienterne, da de beskriver den gennemsnitlige effekt af den uafhængige variabel på den afhængige variabel (LeSage & Pace, 2009, s. 37).

4.3.3 Specifikationstest

Inden SDM- og SDEM-modellerne estimeres, så testes der for om valget af disse rumlige modelspecifikationer er aktuelle, og dette gøres som tidligere beskrevet vha. Moran's I test. Såfremt disse er aktuelle, så vil modellerne estimeres ud fra den nævnte metode foroven.

Dernæst vil estimationsresultaternes robusthed tjekkes overfor valget af specifikation af proxies, kontrolvariable, tidsperiode m.m. Dette gøres med henblik på at tjekke, hvor stabile de opnåede resultater er overfor ændringer i de specifikationsvalg, der er blevet foretaget.

4.4 Databeskrivelse

I dette afsnit præsenteres projektets datagrundlag, som er blevet anvendt til at estimere de økonomiske modeller, som bliver opstillet i kommende afsnit 4.5.

Udgangspunktet for projektet er de europæiske regioner i EU-15 landene, og al regional data blev hentet fra Eurostats database. Eurostat har klassificeret regionale data efter NUTS-klassifikationen (*Nomenclature of Territorial Units for Statistics*), som er en geografisk inddeling af områder. Klassifikationen inddeler de geografiske områder i tre niveauer, herunder NUTS1, NUTS2 og NUTS3. NUTS-klassifikationen er hierarkisk, dvs. NUTS1 inddeler landene i større regioner, mens NUTS2 er en yderligere opdeling af førnævnte niveau, og mens NUTS3 er den mindste inddeling. NUTS-niveauerne er principielt defineret efter populationstørrelse, men favoriserer landenes administrative inddelinger af hensyn til datatilgængelighed. Der tages hovedsageligt udgangspunkt i NUTS2-regioner, men med udtagelse af London. London anvendes på NUTS1-niveau, fordi London i år 2010 gik fra at være opdelt i to NUTS2-regioner, til at være opdelt i fem NUTS2-regioner, hvilket skaber databrud. Dette valg har sine bagsider, da det på NUTS1-niveau ikke er muligt at opfange den store variation der er mellem regionerne på NUTS2-niveau. I 2015 havde den rigeste af disse fem regioner en BNP per indbygger på €167.500, hvorimod den fattigste til sammenligning havde en BNP per indbygger på €22.500 (Eurostat, 2017). På grund af datamangel udelukkes en række regioner af hensyn til at sikre en tilstrækkelig lang tidsperiode, og derudover udelukkes regioner som tidligere nævnt, hvis de ikke har nogen nabo. På baggrund af disse udelukkelser er der blevet konstrueret en stikprøve bestående af $n = 167$ regioner.¹³

Regioner på NUTS2-niveau er blevet udvalgt som geografisk enhed for at sikre tilstrækkelig datatilgængelighed. Af hensyn til at sikre data over en tilstrækkelig lang periode på minimum 10 år, tages der udgangspunkt i regionerne i EU-15 lande, hvorfor der ses bort fra landene i

¹³ For en fuld liste over inkluderede regioner se bilag 1.

østudvidelsen i 2004, og andre lande optaget herefter. En periode på minimum 10 år er at foretrække, da projektet tager udgangspunkt i en teori, der omhandler langsigtet vækst. Der er arbejdet ud fra perioden mellem år 2002 og 2014, men data har ikke været tilgængelig for hele årrækken ved alle regioner, hvorfor der ved datamangel er blevet anvendt data fra årstal tæt på de nævnte år.

4.5 Den økonometriske model

I dette afsnit præsenteres først den økonometriske model, og dernæst præsenteres overvejelserne der ligger bag specificationen af denne. Efterfølgende præsenteres variablene og tilhørende deskriptiv statistik. Afslutningsvis præsenteres overvejelser omkring variable, der kunne være relevante at inkludere, men som ikke er blevet det.

På baggrund af den tidligere præsenterede litteratur om regional vækst og de metodologiske overvejelser, opstilles følgende økonometriske modeller, som vil blive estimeret:

$$dBNP_{cP} = \rho W \cdot dBNP_{cP} + \alpha_n + X\beta + WX\gamma + \varepsilon \quad (\text{SDM})$$

$$dBNP_{cP} = X\beta + WX\gamma + \alpha_n + u \quad (\text{SDEM})$$

Hvor X er vektoren af de uafhængige variable, $X = \{hrst02, tPAT02, dcommu\}$. I Tabel 2 vises hvordan disse modellers variable er blevet operationaliseret.

Tabel 2 Variabeloversigt

Variabel	Forkortelse	Beskrivelse
Vækst i velstand	dBNP	Vækst i BNP per indbygger opgjort i KKP., 2002-2014. Eurostat (2017a).
Tavs viden	hrst02	Humane ressourcer indenfor S&T i pct. af arbejdstyrken, i år 2002. Eurostat (2017a).
Kodificeret viden	tPAT02	Antal ansøgte patenter per mio. indbyggere i år 2002. Eurostat (2017a).
Pendling	dcommu	Udviklingen i antal personer der pendler ud af sin bopælsregion for at arbejde i perioden 2002-2014. Eurostat (2017a) & egne beregninger.

Disse variable er blandt andet valgt med inspiration fra Romer (1990). Ifølge Romer genereres vækst gennem teknologisk udvikling, som i projektet forstås ved produktion af viden, idet hver ny enhed af viden (i relation til teknologi) korresponderer med et design for et nyt gode. Romer antager, at alle nye designs ender med at blive kommercialiseret, hvorfor alle designs også bliver patenteret. Med afsæt i Romers teoriramme forventes det, at jo højere niveau af vidensproduktion der haves i år 2002, desto større udslag vil det give i regioners vækst over perioden 2002-2014. I henhold til kravene for hvordan et patent skal udformes jf. afsnit 4.5, så ligger der implicit et krav om at kodificere sin viden i forbindelse med udarbejdelsen af patentet,

hvorfor antallet af patentansøgninger i år 2002 bliver anvendt som proxy for kodificeret viden. En anden form for viden, som er mindst lige så vigtig for at kunne skabe vækst, er tavs viden. Tavs viden bygger meget på personlig erfaring, og kan i Romers teoriramme tænkes at opstå i forskningssektoren under produktionen af designs, og derfor være manifesteret i humankapitalen. På denne baggrund anvendes antal ansatte i S&T i år 2002, som en proxy for tavs viden. Det er især tavs viden i form af know-how, der forventes at blive fanget, og denne type af viden er også vigtig i forbindelse med at udnytte visse typer af kodificeret viden, som kræver særlige færdigheder og faglige indsigter. Variablen *hrst02* omfatter ikke kun forskere og udviklere, men også alle andre som arbejder i S&T-sektoren. Kodificeret og tavs viden måles begge i vækstperiodens startår, dvs. år 2002, og dette er med henblik på undgå endogenitetsproblemer, som påpeget af LeSage & Fischer (2008, s. 288). Dette problem kan opstå, hvis der både anvendes udviklingen i de afhængige og uafhængige variable over samme tidsperiode, da det ikke kan udelukkes, at kausaliteten i så fald også kan gå den omvendte vej, dvs. at udviklingen i den afhængige variabel kan have en effekt på udviklingen i de uafhængige variable.

Udgangspunktet for Romer er, at ny viden bliver tilgængelig for alle agenter i en økonomi, og at dette afføder positive spillovers af viden indenfor økonomien.¹⁴ Det er jf. den præsenterede teori også muligt, at der sker positive spillovers af viden imellem regioner, men jf. den regionale udviklingslitteratur vil den være i en mindre størrelsesorden, desto større afstande der er mellem regioner, hvorfor geografisk nærhed også er vigtigt for rækkevidden af spillovers og deres effekt. Spillovers af viden mellem naboregioner er derfor modelleret ind i den økonometriske model ved hjælp af et spatial lag af de uafhængige variable, $WX\gamma$, og der tages udgangspunkt i forskellige distancer for at undersøge rækkevidden for spillover-effekter. Ifølge Basile, Capello & Caragliu (2011, s. 25) havde relationel nærhed også en betydning for spillover-effekter, hvorfor der yderligere undersøges, om der er positive spillover-effekter på tværs af landegrænser, idet regioner på tværs af landegrænser formodentligt vil have større kulturelle forskelle end regioner indenfor landegrænser vil.

Der tages også højde for andre regioners vækst, idet $\rho W \cdot dBNPcP$ inddrages i modellen. Baggrunden for dette er, at det formodes at en stigning i velstanden hos andre regioner også vil smitte af på ens egen region. Ifølge den præsenterede

¹⁴ Ny viden bliver tilgængelig for alle, men kan på grund af patenter ikke kommerialiseres af alle.

4.5.1 Variabelbeskrivelse

I nærværende afsnit fremstilles det, hvordan projektet operationaliserer den økonometriske models variable, som blev præsenteret i Tabel 2, og hvilke overvejelser der ligger bag. Efterfølgende præsenteres deskriptiv statistik for de valgte variable i afsnit 4.5.1.1.

BNP per indbygger i KKP (dBNP). Modellens afhængige variabel skal være et mål for væksten i velstand, og som proxy herfor anvendes væksten i BNP per indbygger opgjort i købekraftsparitet (KKP). Opgørelsen i KKP giver en mulighed for bedre at kunne sammenligne regionerne, da der her bliver taget højde for valutakurser og prisforskelle landene imellem (Eurostat, 2015). BNP per indbygger er et meget traditionelt mål for velstand, og anvendes eksempelvis af Rodríguez-Pose & Crescenzi (2008), som tidligere nævnt, men her vælges den i form af KKP grundet førnævnte årsag. En ulempe ved at bruge dette mål for velstand er, at personer ikke nødvendigvis bor i den region, hvor de skaber værditilførslen, idet de pendler på tværs af regionsgrænser for at arbejde.

Totale antal patentansøgninger (tPAT02). Totale antal patentansøgninger bruges som en proxy for hvor meget kodificeret viden der er produceret i en region i år 2002. Baggrunden for dette er, at der i ansøgningsprocessen hos det europæiske patentkontor (EPO) er et krav om, at vedlægge en udførlig beskrivelse af opfindelsen og dens design, således at en person med samme faglige forudsætninger kan replikere designet (European Patent Office, 2016, s. 25). Processen kræver således, at viden bliver kodificeret, hvorfor patentansøgninger kan bruges som en proxy for kodificeret viden. Kodificeret viden angives således i antal ansøgte patenter til EPO per millioner indbyggere i år 2002, altså i startåret for vækstperioden, og ikke i udviklingen i antal ansøgte patenter over vækstperioden. Dette skyldes, at det tager tid før patenter udnyttes og kommerialiseres, og derved har en effekt på væksten i den observerede periode. Ydermere er det for at undgå endogenitetsproblemer, som tidligere påpeget. Det forventes som sagt, at variabelen vil have en positiv effekt på væksten, både direkte og indirekte (spillovers), da det forventes at denne form for viden har nemmere ved at rejse på tværs af regionsgrænser.

Fordelen ved at anvende totale patentansøgninger er, at de dækker et stort område indenfor mange forskellige teknologier og brancher. Derudover er tilgængeligheden af patentdata meget stor, og det giver en god indikation på, hvordan udviklingen i en region skrider fremad. Denne proxy har dog også sine begrænsninger, da det ikke er alle patenter der bliver til et kommercielt produkt. Det anslås at ca. 40 pct. af tildelte patenter aldrig kommer i produktion, enten fordi

patentet er et forsøg på at blokere konkurrenter, eller også bliver patenterne aldrig brugt, fordi de ikke kan opnå en tilstrækkelig kommerciel værdi for opfinderen (OECD, 2009c, s. 26).

Bagsiden ved, at proxien dækker alle brancher er, at den ikke informerer om variationen i antal patentansøgninger på tværs af brancherne. Visse brancher har en højere tilbøjelighed til at søge patenter end andre, såsom medicinal- og biotekbranchen, som er velkendte patentintensive industrier, hvorimod serviceindustrier ikke har megen tilbøjelighed til at søge patenter (OHIM & EPO, 2013, s. 9). Implikationen af dette er, at regioner der primært består af brancher, som har en lav tilbøjelighed til at søge patenter, kan fremstå som mindre vidensproducerende end de rent faktisk er.

Antal ansatte indenfor science og technology (hrst02). Antal ansatte indenfor science and technology (S&T) bruges som en proxy for tavs viden. Viden i forbindelse med S&T er oftest tavs, da det er en type viden, der læres gennem mange års erhvervserfaring (Bilbao-Osorio & Rodríguez-Pose, 2004, s. 438). Tavs viden opbygges bl.a. i denne sektor fordi denne type ansatte beskæftiger sig med F&U-relaterede aktiviteter i processen med at generere designs. Definitionen af ansatte i S&T er, at de skal have fuldført en tertiær uddannelse, eller være ansat indenfor branchen i en stilling, der normalt kræver en tertiær uddannelse. Antal ansatte indenfor S&T er opgjort i pct. af arbejdsstyrken, som i Eurostat er defineret af folk i aldersgruppen 15-74-årige, som er tilgængelige for arbejdsmarkedet. Der tages udgangspunkt i antal ansatte i S&T i år 2002. HRST02 forventes at fange i tavs viden i form af blandt andet know-how, og den fanges idet ansatte igennem uddannelse og erhvervserfaring tilegner sig færdigheder, der gør dem i stand til at udnytte og arbejde med den eksisterende viden. Det forventes at variabelen vil have en positiv direkte effekt på væksten, mens den indirekte effekt er tvetydig, idet tavs viden har sværere ved at rejse over længere distancer, hvorfor den muligvis ikke har nogen effekt, men såfremt der er en effekt, forventes den at være positiv.

Fordelen ved at bruge denne variabel som proxy er, at der fokuseres på et bestemt segment indenfor tavs viden, som er det segment, hvor der er et relativt højt uddannelsesniveau, da langt størstedelen har en tertiær uddannelse. Tertiær uddannelse kunne være en alternativ proxy, men denne størrelse vil tilbyde en snæver forståelse af tavs viden, da den ikke inkluderer tavs viden opnået igennem erhvervserfaring. I henhold til Romers definition på humankapital, så fanger hrst02 at færdigheder tilegnes ud fra både uddannelse og erhvervserfaring, i modsætning til tertiær uddannelse, der kun fanger det uddannelsesmæssige aspekt, hvorfor hrst02 anvendes som en proxy for tavs viden.

Pendling (dcommu). Udviklingen i antal pendlere er medtaget, grundet svagheden ved at bruge BNP per indbygger som proxy for vækst i velstand. Monfort (2009) refererer til en *pendlingseffekt*, der har en tendens til at overestimere BNP per indbygger i store byer, mens den underestimerer BNP i det omkringliggende pendlingsopland (Monfort, 2009, s. 6). Således kontrolleres der for, at en person ikke nødvendigvis bor og arbejder i samme region. Målet er at se, om det har en betydning for væksten, hvor mange flere personer der pendler ud af regionen for at arbejde. Det forventes at have en negativ direkte effekt på regional vækst, hvis der er flere der pendler ud af regionen, da det må tyde på, at der genereres mindre velstand indenfor regionen, mens den indirekte effekt forventes at være positiv, da det kan være et udtryk for, at en andel af dem der pendler ud af naboregionen j , sandsynligvis pendler ind i region i . Pendling opgøres i antal tusinde, der arbejder udenfor den region de har bopæl i.

4.5.1.1 Deskriptiv statistik

I følgende afsnit vil variablene beskrives ved hjælp af deskriptiv statistik, og de vil ydermere præsenteres visuelt. Indledningsvist beskrives variationen mellem regionerne inden for modellens variable. Derefter kommer en mere detaljeret gennemgang af variablene. Tabel 3 giver en oversigt over standardafvigelsen, middelværdien, samt maksimum og minimum for variablene.

Tabel 3 Deskriptiv statistik

Variable	Middelværdi	Standardafvigelse	Minimum	Maksimum
dBNP	20,743	0,09942893	-0,9	47,589
hrst02	34,47	6,767491	11,50	53,60
tPAT02	123,459	128,2254	1,249	829,405
dcommu	20,785	0,4611831	-233,856	210,261

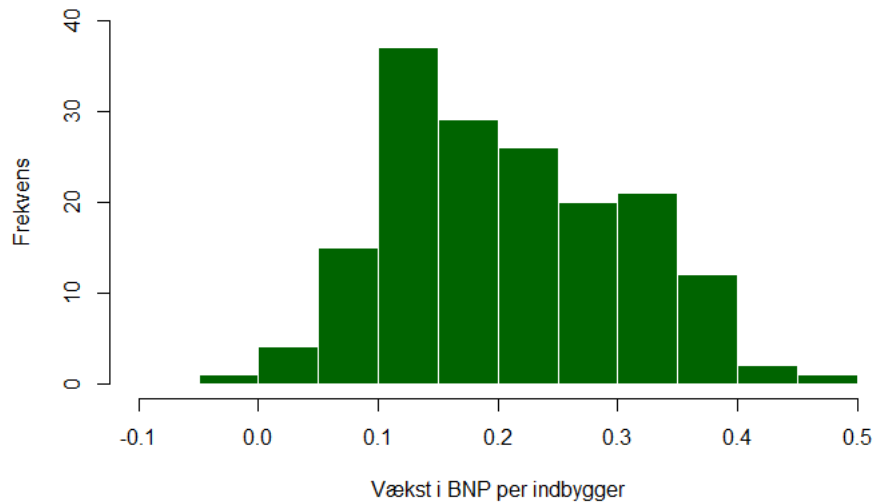
Note: dBNP, hrst02 og dcommu er angivet i pct.

Kilde: Egne beregninger, baseret på Eurostat (2017a)

BNP per indbygger i KKP (dBNP).

Det fremgår af Tabel 3, at BNP per indbygger i regionerne i gennemsnit er steget med 20,7 pct. over vækstperioden fra 2002–2014. Her er det vigtigt at bemærke, at væksten er over en periode på 12 år, og er derfor ikke angivet i årlige vækstrater, hvorfor det er muligt at opnå en vækst på 47,6 pct. Ses der på fordelingen af den regionale vækst, så viser Figur 5, at de fleste regioner har en vækst mellem 10 og 20 pct. Det ses samtidig også, at maksimum- og minimumsobservationen begge er de eneste indenfor deres interval. Generelt viser Figur 5, at alle regioner, undtagen én, har oplevet en øget velstand i perioden 2002-2014.

Figur 5 Histogram over væksten i BNP per indbygger i perioden 2002–2014.



Kilde: Egne beregninger, baseret på Eurostat (2017a)

Grundet projektets fokus på geografiske forskelle, er det nærliggende at se den geografiske placering af de regioner, der har haft en høj eller lav vækst. I Tabel 4 præsenteres derfor en top ti og en bund 10 over regioner med højeste og laveste vækstrater.

Tabel 4 Top 10 og bund 10 for vækst i BNP per indbygger.

Top 10		Bund 10	
Braunschweig (DE)	47,59 pct.	Franche-Comté (FR)	-0,9 pct.
Thüringen (DE)	42,63 pct.	West Midlands (UK)	2,13 pct.
Sachsen-Anhalt (DE)	40,13 pct.	Umbria (IT)	3,39 pct.
Brandenburg (DE)	38,71 pct.	Balearic Islands (ES)	3,89 pct.
Dresden (DE)	38,45 pct.	Campania (IT)	4,28 pct.
Tubingen (DE)	38,42 pct.	Lazio (IT)	5,22 pct.
Mecklenburg-Vorpommern (DE)	38,41 pct.	Bedfordshire and Hertfordshire (UK)	6,00 pct.
Arnsberg (DE)	37,53 pct.	Piedmont (IT)	6,36 pct.
Wesser-Ems (DE)	37,47 pct.	Friuli-Venezia Giulia (IT)	7,08 pct.
Oberfranken (DE)	37,09 pct.	Marche (IT)	7,38 pct.

Note: Landekode er angivet i ().

Kilde: Eurosta (2017a) t

Her er det værd at bemærke, at de ti regioner der har oplevet den højeste vækst, alle er tyske regioner. Indenfor top 10 er der en relativ stor spredning mellem regionerne. Braunschweig, som ligger nr. 1, har haft en vækst der er 10,5 pct. point højere end nr. 10, Oberfranken.

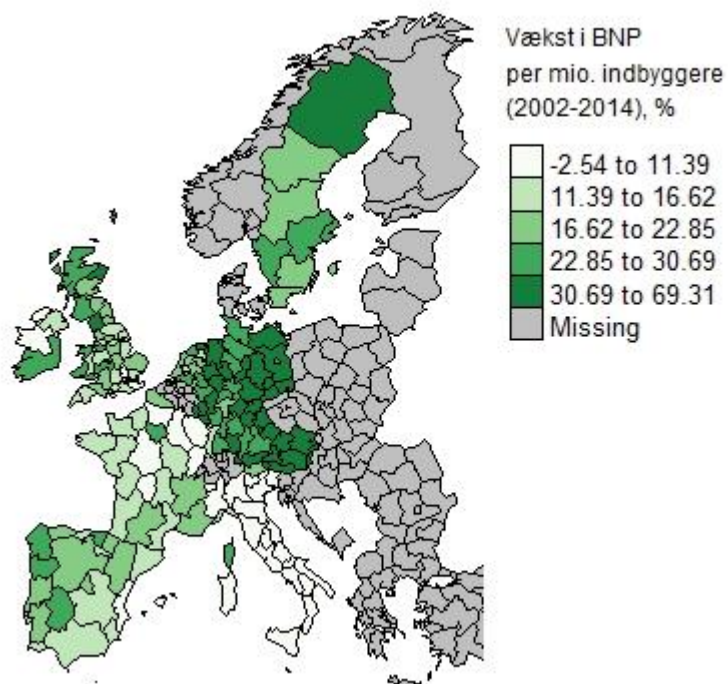
Braunschweigs førerposition i forhold til vækst i BNP per indbygger kan bl.a. skyldes, at Braunschweig i 2005 var den region i EU, hvor F&U-investeringer udgjorde den største andel af BNP (Petkova, 2009). Braunschweig har i relation til F&U-investeringer ligget i top i de

efterfølgende år. Derudover udgøres 6 ud af de 10 regioner med laveste vækst af italienske regioner. Den franske region Franche-Cometé er den eneste region, der har oplevet et fald i velstanden. Dette er dog misvisende. Franche-Cometé er en industriregion med en meget specialiseret industrisammensætning, og er den region i Frankrig, der har den største eksport (Europa-Kommissionen, 2017a). Eksempelvis produceres Peugeot-Citroën i denne region, og de er derfor meget følsomme overfor konjunktursvingninger, hvorfor det har stor betydning, hvilken tidsperiode væksten observeres over.

For at kunne få et indtryk af hvordan høj- og lavvækstregioner fordeler sig geografisk, så vil der ud fra europakortet i Figur 6 laves en visuel inspektion. Dette er med henblik på at se, om der er en tendens til at højvækstregioner udgør naboer, og om lavvækstregioner udgør hinandens naboer. Figur 6 er som sagt et kort over de europæiske regioner, og kortet viser væksten i BNP per mio. indbyggere i perioden 2002-2014. Mørkere farver markerer, at regionerne har en relativ større vækst i BNP per mio. indbyggere, mens lysere regioner har en relativt mindre vækst i BNP per mio. indbyggere.¹⁵ Grænserne mellem farverne er defineret efter datas kvartiler, således at den første grænse er intervallet mellem minimumsværdien og det første kvartil osv. Det fremgår af Figur 6, at der er en stor geografisk koncentration af højvækstregioner i Tyskland, som indikeret af de mørkegrønne farver, men også de østrigske ligger i det øverste interval. Samtidig ses der en geografisk koncentration af lavvækstregioner i Italien, som indikeret af de lyse farver, hvilket ikke er overraskende, da de italienske regioner også var stærkt repræsenteret i bund 10 i Tabel 4. Derudover har de franske regioner, der omringer Paris-regionen, også haft lav vækst i perioden.

¹⁵ Bemærk, at kortet i Figur 6 ligesom med Figur 8 og Figur 9 er baseret på de nyeste data fra eurostat, hvorfor der forekommer afvigelser mellem max og min angivet i kortet samt de angivet i Tabel 4, Tabel 5 og Tabel 6.

Figur 6 Vækst i BNP per mio. indbyggere i perioden 2002-2014, %.



Note: Grå markerer enten manglende data for perioden, eller at regionen ikke indgår i stikprøven.

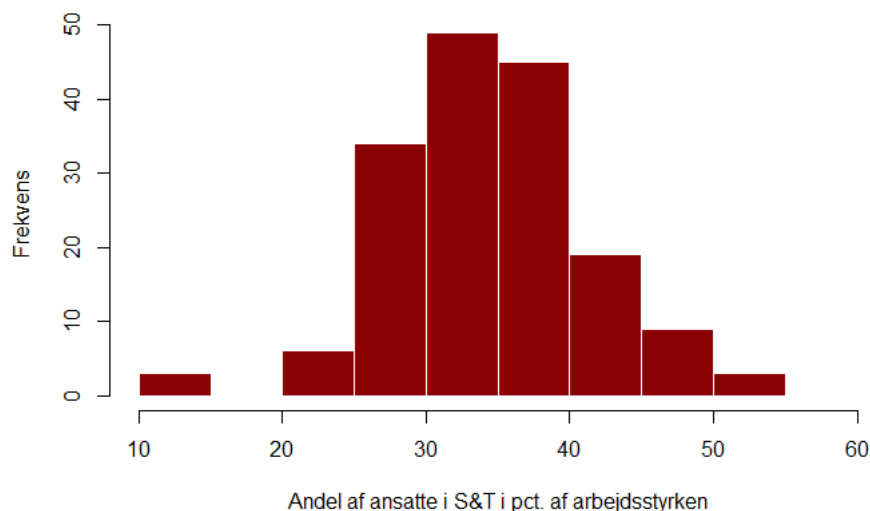
Kilde: Egne beregninger på baggrund af data fra (Eurostat, 2017a).

© EuroGeographics for de administrative grænser

Antal ansatte indenfor science og technology (hrst02).

Fokuseres der i stedet på antal af ansatte i S&T i pct. af arbejdsstyrken (hrst02), så fremgik det af Tabel 3 at denne type ansatte i gennemsnit udgør 34,37 pct. af arbejdsstyrken i de europæiske regioner, og fordelingen af hrst02 fremgår af Figur 7.

Figur 7 Histogram over antal af ansatte i S&T i pct. af arbejdsstyrken i 2002.



Kilde: Egne beregninger, baseret på Eurostat (2017a)

Figuren viser at ansatte i S&T udgør en relativ stor del af arbejdsstyrken i langt de fleste regioner. Med undtagelse af tre regioner, udgør antal ansatte i S&T 20 pct. eller derover af arbejdsstyrken. Top 3 adskiller sig dog fra resten, idet over 50 pct. af deres arbejdsstyrke er beskæftiget i S&T. Ligesom tidligere er der nu lavet en top 10 og bund 10, men denne gang for antal ansatte i S&T, og denne er præsenteret i Tabel 5 forned.

Tabel 5 Top 10 og bund 10 for antal ansatte i S&T i pct. af arbejdsstyrken i 2002

Top 10		Bund 10	
Stockholm (SE)	53,6 pct.	Central Portugal (PT)	11,5 pct.
Valloon Brabant (BE)	52,1 pct.	North (PT)	12,8 pct.
Utrecht (NL)	50,4 pct.	Alentejo (PT)	15,0 pct.
Ile-deFrance (FR)	49,0 pct.	Basilicata (IT)	22,9 pct.
Bruxelles-Cap. (BE)	48,9 pct.	Aosta valley (IT)	23,3 pct.
Noord-Holland NL)	47,2 pct.	Calabria (IT)	24,5 pct.
Berlin (DE)	47,0 pct.	Province of Bolzano-Bozen (IT)	24,7 pct.
London (UK)	46,5 pct.	Apulia (IT)	24,9 pct.
Madrid (ES)	45,9 pct.	Extremadura (ES)	24,9 pct.
Hamburg (DE)	45,5 pct.	Castile-La Mancha (ES)	25,4 pct.

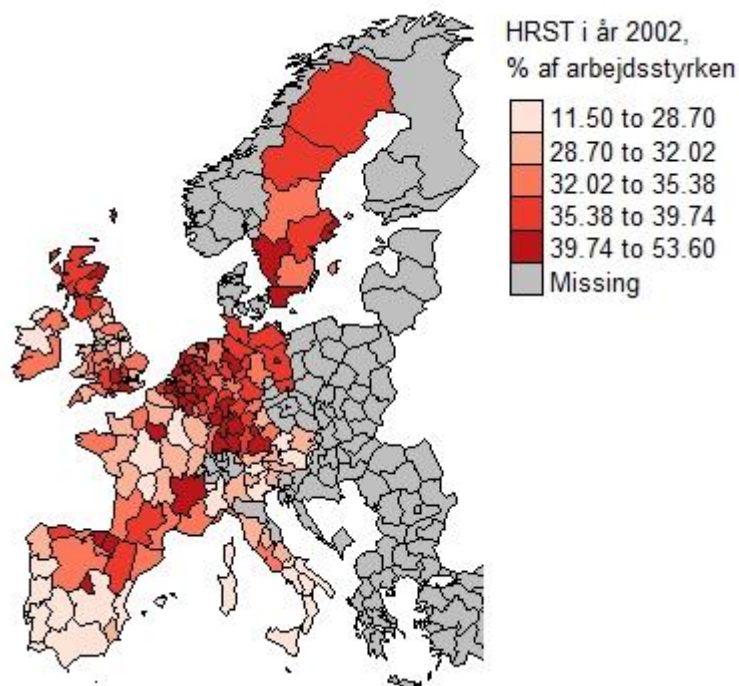
Kilde: Eurostat (2017a)

Her er det interessant at bemærke, at syv ud af ti af de regioner, hvor antal ansatte i S&T udgør den største del af arbejdsstyrken, er hovedstadsregioner. Igen er italienske regioner velrepræsenteret i bund 10, idet halvdelen udgøres af italienske regioner, men de er dog ikke de regioner med det laveste antal ansatte i S&T i pct. af arbejdsstyrken. De tre regioner der har den mindste andel af deres arbejdsstyrke i S&T, er alle portugisiske regioner, og der er et spring på

7,9 pct. point mellem nr. tre og fire. Disse tre skiller sig ud fra resten, hvilket også fremgik af histogrammet i Figur 7. S&T-sektoren er relativ nystartet i Portugal, og den er derfor stadig under udvikling. En af grundene til, at de portugisiske regioner præsterer så dårligt i dette henseende kan bl.a. skyldes det dårlige uddannelsessystem i Portugal. Det er kun for ganske nyligt, at Portugal selv har været i stand til at udbyde de uddannelser det oftest kræver for at arbejde i S&T, og førhen måtte portugisere rejse til udlandet for at tage de rette uddannelser (Coutinho, 2013) .

At hovedstadsregionerne er stærkt repræsenteret i top ti, fremgår også af europakortet i Figur 8. I hovedstadsregionerne og regionerne i Benelux-landene og det sydlige Tyskland udgøres arbejdsstyrken af en relativ stor andel af S&T ansatte. Regionerne i Sydeuropa er igen, sammen med regionerne omkring Paris, velrepræsenteret i det laveste interval, som markeret af de lyse farver.

Figur 8 Antal ansatte indenfor S&T i år 2002, i pct. af arbejdsstyrken.

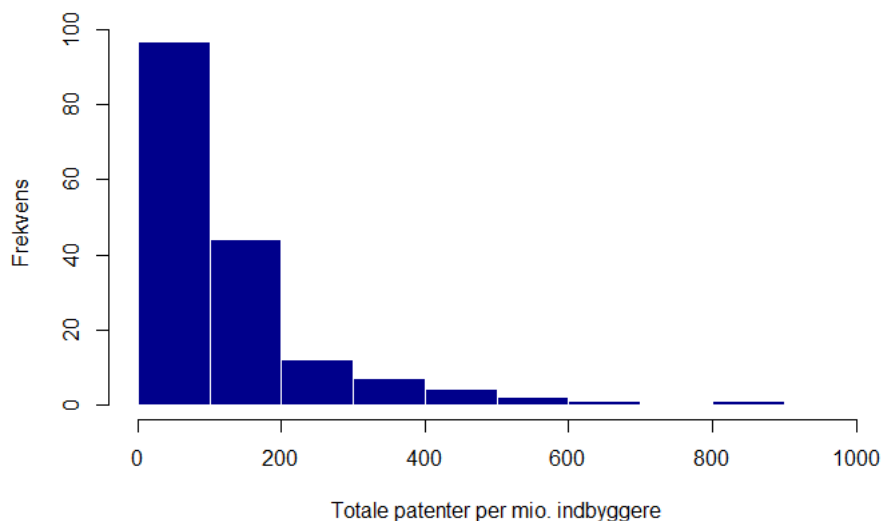


Note: Grå markerer enten manglende data for året, eller at regionen ikke indgår i stikprøven.
Kilde: Eurostat (2017a), © EuroGeographics for de administrative grænser.

Totale antal patentansøgninger (tPAT02).

Gennemsnittet for de europæiske regioners patentansøgninger er lig 123,459, og dette hænger sammen med, at langt de fleste regioners antal af patentansøgninger ligger mellem 0 og 200, som det fremgår af Figur 9.

Figur 9 Histogram over totale antal patentansøgninger per mio. indbyggere i 2002.



Kilde: Egne beregninger, baseret på Eurostat (2017a).

Selvom langt de fleste regioner har en patentaktivitet på under 100 i 2002, er de få regioner, der har en meget stor patentaktivitet, i stand til at trække gennemsnittet op. Laves en top 10 og bund 10, så fås følgende:

Tabel 6 Top 10 og bund 10 for antal ansøgte patenter per mio. indbyggere i 2002

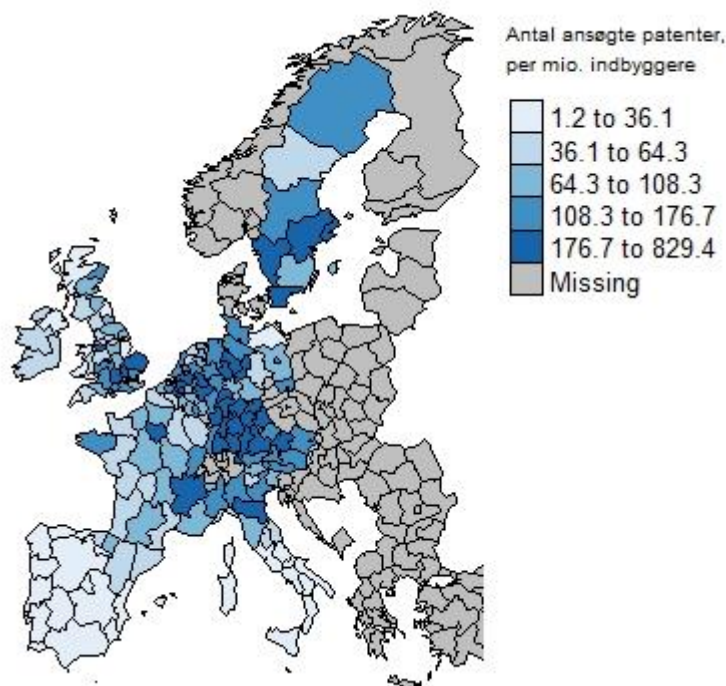
Top 10		Bund 10	
Noord-Brabant (NL)	829,405	Cantabria (ES)	1,249
Stuttgart (DE)	668,631	Alentejo (PT)	1,610
Oberbayern (DE)	568,234	Extremadura (ES)	1,891
Tubingen (DE)	501,496	Molise (IT)	3,124
Mittelfranken (DE)	480,951	Balearic islands (ES)	3,349
Karlsruhe (DE)	473,825	Central Portugal (PT)	3,402
Rheinhessen-Pfalz (DE)	445,003	Lisbon (PT)	4,271
Freiburg (DE)	433,957	North (PT)	4,688
Darmstadt (DE)	389,823	Castile-La Mancha (ES)	5,172
Unterfranken (DE)	380,003	Basilicata (IT)	5,577

Kilde: Eurostat (2017a)

Heraf fremgår det, at de tyske regioner igen dominerer i top 10, men de formår at blive overgået i antal patentansøgninger af den hollandske region, Noord-Brabant. Spredningen indenfor top 10 er meget stor, da forskellen mellem nr. et og nr. ti er ca. 450 patenter per mio. indbyggere. Selv forskellen mellem nr. et og nr. to er relativt stor. Den hollandske region er karakteriseret ved at have en særlig industrisammensætning, idet den primært består af hightech, biovidenskab og sundhed, som tilsammen står for 76 pct. af den totale produktion i regionen. Den hollandske by Eindhoven har en stor andel i regionens succes, og byen har fået tilnavnet ”brainport” pga. en meget stor koncentration af hightech virksomheder, og fordi den står for 44 pct. af de samlede patentansøgninger fra hele Holland (Europa-Kommisionen, 2017b). Bund 10 er igen domineret af sydeuropæiske lande, som har et relativt lavt antal af patentansøgninger, hvilket vidner om mindre vidensproduktion.

De tyske regioners dominans fremgår også af europakortet i Figur 9.

Figur 9 Antal ansøgte patenter per mio. indbyggere, år 2002.



Note: Grå markerer enten manglende data for året, eller at regionen ikke indgår i stikprøven.
 Kilde: Eurostat (2017a), © EuroGeographics for de administrative grænser.

Her er det tydeligt, at især de sydtyske regioner har mange patentansøgninger, og at disse regioner er geografisk koncentreret. Disse sydtyske regioner rummer mange af de samme hightech-industrier, og disse industrier har en højere tilbøjelighed til at søge patenter. En af disse

sydtyske regioner er Stuttgart, som rummer mange internationale virksomheder, som inkluderer IBM, Bosch, Porsche og Daimler som laver Mercedes-Benz (Europa-Kommissionen, 2016). En nærliggende region, Oberbayern, har lignende virksomheder, og disse omfatter bl.a. BMW og Siemens (Europa-Kommissionen, 2004). Denne geografiske koncentration af tyske regioner, der har en høj vidensproduktion, kunne på den ene side være tilfældig, men på den anden side kunne denne koncentration vidne om, at regionerne drager nytte af at være tæt placeret på regioner med lignende industrispecifik viden. Det kunne vidne om, at regionerne har en fordel ved at være placeret på regioner med hvem de har en kort kognitiv distance jf. Boschma (2005). Derudover er de belgiske og hollandske regioner, som sagt også stærkt repræsenteret i forhold til antal patentansøgninger. Igen ses der en kontrast mellem Paris-regionen og dens omkringliggende regioner, men forskellen er ikke lige så stor som den var ved de forrige variable. En anden tendens der går igen, er den manglede aktivitet i Spanien, Portugal og Italien, men denne gang eksisterer der et tydeligt skel mellem Nord- og Syditalien. Norditaliens regioner har væsentlig flere patentansøgninger end Syditalien har. Dette kan forklares ud fra den observation, at en af dens regioner, Emilia-Romagna, har en meget stor koncentration af medium-høj tech-industrier, som omfatter jern- og metalindustrier, samt en bilindustri hvor der produceres Ferrari, Maserati og Lamborghini (Europa-Kommissionen, 2017c).

Opsummering af deskriptiv statistik

Dette afsnit viser, at de tyske regioner dominerer top 10 både i relation til vækst i BNP per indbygger og i antal ansøgte patenter. For antal ansatte i S&T i pct. af arbejdsstyrken var der dog en stor overvægt af hovedstadsregioner, men de har også selskab af sydtyske regioner, som ligger indenfor det højeste interval. Europakortene viser at stærke regioner, i form af de tyske regioner, og her især de sydtyske, har en tendens til at være geografiske naboer. En undtagelse er Paris-regionen og dens omkringliggende regioner, hvorimellem kontrasten er stor, idet de ikke deler mange karakteristika. Det ligner at Paris fungerer som en støvsuger for naboregionerne.

I bund 10 var der for alle variable en meget stor overvægt af sydeuropæiske regioner. Det var især de italienske og portugisiske regioner, som var at finde i bund 10. Der er således en tendens til en nord/syd opdeling af de europæiske regioner i EU-15. Det fremgår også af europakortene, at der er en tendens til at regioner, der klarer sig godt i relation til hver variabel, også i de fleste

tilfælde har en stærk region som nabo, og det modsatte gør sig også gældende ved regioner, der klarer sig mindre godt i de forskellige henseender.

4.5.2 Variable til robusthedstjek

I afsnit 5.2 præsenteres resultaterne af den økonometriske analyse på baggrund af de ovennævnte variable, og dernæst foretages en robusthedstjek, hvor der kontrolleres for udeladte variable, som af regional vækstlitteraturen anses for at være vigtige for vækst. Disse variable inkluderer agglomeration og en hovedstadsdummy. Ydermere vil der kontrolleres for om udviklingsniveau har betydning for vækstpræstationen i den observerede periode.

Agglomeration (densi). Befolkningstætheden, dvs. antal personer per km², anvendes som proxy for agglomeration. Med denne variable forventes det at kunne opfange de eksternaliteter, der kan være til stede ved en større koncentration af mennesker i regionen. Agglomeration kan på den ene side være et udtryk for, at der eksisterer positive eksternaliteter ved at befinde sig i en større koncentration af mennesker. Viden og ikke mindst den tavse af slagsen, vil lettere kunne overføres mellem personer og virksomheder. LeSage & Fischer (2008, s. 290) inkluderer også denne variabel i deres analyse af regional vækst, idet de argumenterer, at regioner der indeholder en større befolkningstæthed, repræsenterer agglomerationer, som indeholder en større beholdning af viden lagret i humankapitalen, hvilket er med til at fremme teknologisk udvikling og derved vækst. På den anden side kan der som tidligere nævnt ikke udelukkes, at der kan være negative eksternaliteter til stede ved en større agglomeration. Disse eksternaliteter vil, alt afhængig af hvilken der dominerer, manifesterer sig i højere eller lavere vækst. F.eks. vil en dominans af positive eksternaliteter forventes at have en positiv effekt på vækst.

Hovedstadsdummy (capital). Ud fra den visuelle inspektion af Figur 8 fremgår det, at en stor del af hovedstadsregioners arbejdsstyrke er ansat i S&T. Dette bekræftes også i Tabel 5, da de dominerer top-10. Derfor kontrolleres der for betydningen af at være en hovedstadsregion, altså en region med landets hovedstad. Dette operationaliseres ved hjælp af en dummyvariabel, der tager værdien 1, hvis regionen indeholder landets hovedstad, og 0 hvis den ikke gør. Cuaresma, Doppelhofer, & Feldkircher (2012, s. 46) inddrager også denne kontrolvariabel i deres studie, og de finder at hovedstadsregioner i gennemsnit vokser mere end regioner uden en hovedstad. På den anden side er der indikation på, jf. senere afsnit, at hovedstadsregionerne i EU-15 i forvejen er velstående regioner, hvorfor regionerne ikke vil opleve de største vækstrater.

Udviklingsniveau. De europæiske regioner i stikprøven har naturligvis forskellige karakteristika, som bidrager til, at de har forskellige forudsætninger for at vækste. Traditionen tro vil der i dette studie også kontrolleres for en regions udviklingsniveau i vækstperiodens startår, og som proxy for udviklingsniveauet anvendes BNP per indbygger i år 2002 (Rodríguez-Pose & Crescenzi, 2008, s. 56; OECD, 2009a, s. 80). I neoklassisk vækstlitteratur har argumentet været, at relativt tilbagestående regioner er i gang med at indhente mere udviklede regioner, hvorfor deres vækstrater må forventes at være højere (Verspagen, 1991, s. 360). Proxyen bruges i nærværende tekst ikke som en uafhængig variabel, men den bruges i stedet til at dele stikprøven op i to delstikprøver. Den første omfatter regioner, som befinder sig i de 25 pct. af observationerne, der ligger over det tredje kvartil (top-25), mens den anden omfatter de der befinder sig i de 75 pct. af observationerne, som ligger under det tredje kvartil (bund-75). Dette giver i udgangspunktet to stikprøver, top-25 og bund-75, der har hhv. $n = 42$ og $n = 125$ regioner, og en gennemsnitlig BNP per indbygger lig hhv. €31.971 og €20.861.¹⁶ Stikprøverne får således vidt forskellige størrelser, men formålet med den tredje kvartil som skillelinje er, at udskille de mest udviklede regioner fra de mindst udviklede, og ligeledes at udskille de mest udviklede fra dem der er på vej til at blive en af de mest veludviklede regioner. Ydermere kan opdelingen af stikprøven gøre det mere tydeligt, om kodificeret viden og tavs viden har lige stor betydning for vækst i relativt veludviklede regioner i forhold til relativt mindre udviklede regioner.

4.5.2.1 Deskriptiv statistik

Som tidligere nævnt vil en relativt tilbagestående region kunne opleve høje vækstrater, hvorfor en regions udviklingsniveau har betydning for hvor store vækstrater en region vil opleve. Derfor kan en top-10 og bund-10 give et overblik over, hvilke regioner der vil opleve de største vækstrater, og hvilke regioner som måske ikke vil opleve helt så store vækstrater, idet de i forvejen er relativ velstående.

¹⁶ Visse regioner er blevet smidt ud under konstruktionen af vægtmatricerne for regressionen med delstikprøverne, idet de manglede naboregioner, og derfor ender de benyttede stikprøver med at have $n = 36/38$ og $n = 124$ regioner.

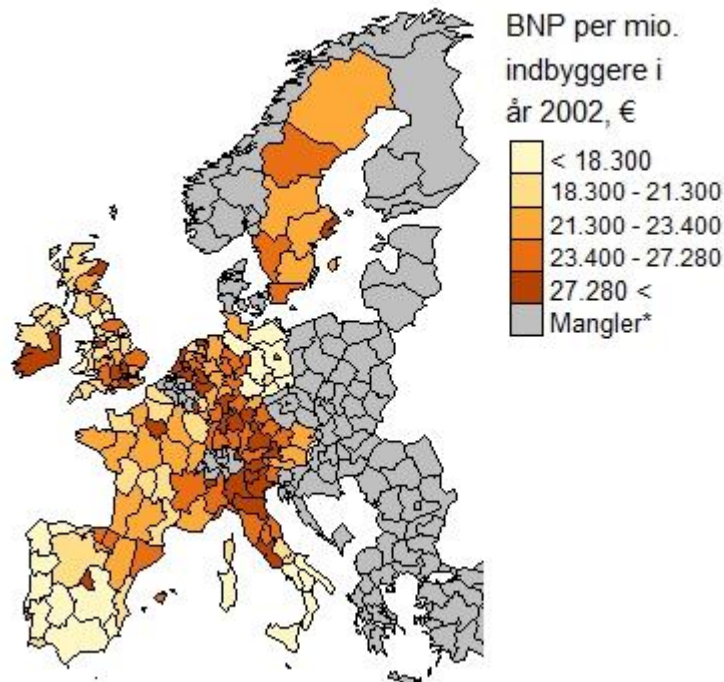
Top-10		Bund-10	
Bruxelle-Cap. (BE)	€52000	North (PT)	€13300
Luxembourg (LU)	€50700	Extremadura (ES)	€13500
Hamburg (DE)	€44400	Central Portugal (PT)	€14000
London (UK)	€38800	Calabria (IT)	€14300
Ile-de-France (FR)	€37900	Alentejo (PT)	€15000
Vienna (AT)	€37300	Sicily (IT)	€15400
Oberbayern (DE)	€36600	West Wales and the Valleys (UK)	€15400
Berkshire, Buckinghamshire and Oxfordshire (UK)	€36600	Apulia (IT)	€15500
Noord-Holland (NL)	€36500	Mecklenburg-Vorpommern (DE)	€15800
Utrecht (NL)	€36300	Thuringen (DE)	€15800

Kilde: Eurostat (2017a)

Tabel 7 viser igen at der i top-10 er en dominans af hovedstadsregioner, hvilket igen giver en grund til at inddrage hovedstadsdummyen. Samtidig er der en stor spredning i top-10, hvor forskellen mellem et og ti er 15700 €. I bund-10 er der igen en overvægt af sydeuropæiske regioner.

I Figur 10 fremgår det af placeringen af de regioner der har det højeste udviklingsniveau, set bort fra de højtudviklede hovedstadsregioner, at der igen er en tendens til at regionerne ligger i små grupperinger. Her er det værd at lægge mærke til, at de norditalienske regioner ligger i den øverste kvartil, og der har ellers været en tendens til at de italienske regioner ikke har været lige så meget repræsenteret i den øverste kvartil ved de andre variable.

Figur 10 Udviklingsniveau i år 2002, målt på BNP per mio. indbyggere i €.



Note: * Markerer enten manglende data for året, eller at regionen ikke indgår i stikprøven.
Kilde: Eurostat (2017a), © EuroGeographics for de administrative grænser.

4.5.3 Alternative variable

Der har ved specifikationen af den økonometriske model været overvejelser om at inkludere andre variable, som også har betydning for vækst.

Romer (1990) fokuserede især på investeringer i F&U som en vigtig drivkraft bag produktionen af viden og derved skabelsen af vækst. Derfor ville det have været nærliggende at inddrage udgifter brugt på F&U som en variabel, der kontrollerer for regionens indsats på vidensudvikling. Årsagen til, at dette ikke er tilfældet er, at F&U-aktiviteter forløber patenter, hvorfor der eksisterer et tidslag mellem disse to størrelser, der ikke berettiger inkluderingen af observationer for F&U-udgifter og patenter i samme tids- og vækstperiode, og det har ydermere ikke været muligt at finde data for F&U før år 2000, for at kunne konstruere en lagget værdi af F&U-udgifter. F&U-udgifter kan anses for at udgøre et input i vidensproduktion, mens patenter derimod kan anses for at være outputtet. F&U-udgifter er i dette henseende ikke nogen garanti for, at der skabes ny viden, og hvis der gør, vil der som sagt gå nogle år.

En af de traditionelle måder at måle humankapital på er ved at bruge tertiær uddannelse som proxy herfor. Fordi uddannelse er relativt nemt at måle og giver en simpel måde at sammenligne uddannelsesniveaet mellem regioner. En af ulemperne er dog at kvaliteten ikke nødvendigvis

er ens i regionerne. Tertiær uddannelse ville her kunne være et alternativ til at måle tavs viden, idet de færdigheder der opnås gennem en uddannelse giver en større tavs know-how. Den kunne også indgå som en kontrol variabel for en regions absorptive evner, men da den kan tænkes at korrelere meget med HRST kan det skabe bias i estimerne. Testes der for korrelation mellem variablene med en VIF-test viser den dog ikke at der skulle være en høj korrelation mellem variablene, men da HRST udgøres af mange personer med tertiær uddannelse vil HRST være en delmængde af tertiær uddannelse, hvorfor de fanger meget af det samme.

I Romer 1990 er kapital også et input i produktionsfunktionen. Selvom Romer identificerer kapital som en vigtig del af vækstprocessen, så modelleres den ikke ind i den økonometriske model, dels fordi der ikke er data for de spanske regioner, dels fordi dette ikke er en empirisk test af Romers teori.

5 Resultater

I nærværende afsnit vil projektets resultater præsenteres. Først testes der for rumlig afhængighed, derefter præsenteres estimationsresultaterne for SDM og SDEM. Herefter præsenteres resultaterne for betydningen af landegrænser for spillover-effekten af viden på regioners vækst. Afslutningsvist vil resultaterne af robusthedstjekkene præsenteres.

5.1 Test for rumlig afhængighed

Indledningsvist er der blevet testet for rumlighed afhængighed mellem residualerne i den lineære model. Til dette formål er der blevet konstrueret en vægtmatrice baseret på et distancebånd, og denne konstruktion har resulteret i en oversigt over naboskab for de to distancebånd, som rapporteret i Tabel 8.

Tabel 8 Oversigt over naboskaber defineret ud fra vægtmatrice

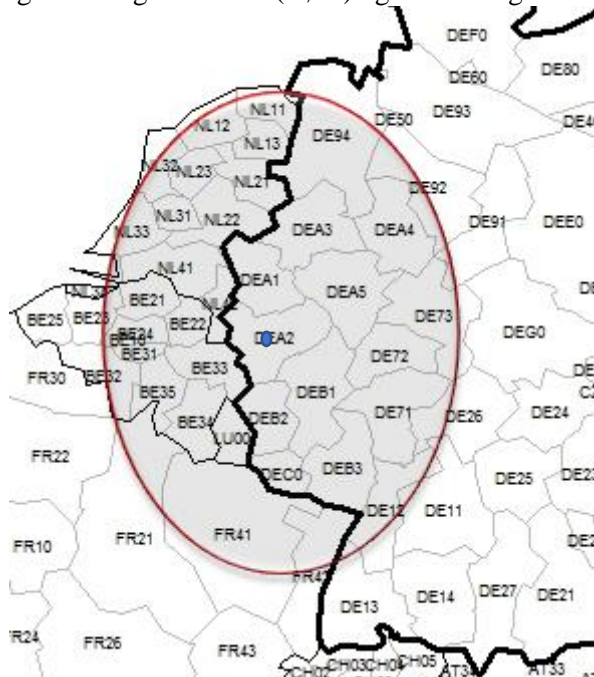
Bånd	Antal naboskaber	Naboskaber i gennemsnit	[Min, Max]
$d_1 \approx]0: 205 \text{ km}[$	2584	15,47	[1,37]
$d_2 \approx]205: 340 \text{ km}[$	2914	17,45	[2,41]

Note: Baseret på $n = 167$ regioner.

Kilde: Egne beregninger.

Ifølge Tabel 8 er der stor forskel på antallet af naboskaber. Indenfor det første bånd er der to regioner der begge har 37 naboer, hvilket er det højeste antal naboskaber en region har i stikprøven. En af disse regioner er Köln. I Figur 11 er Kölns naboerregioner illustreret visuelt.

Figur 11 Region DEA2 (Køln) og dens mange naboer



Note: Dette er ikke en eksakt afbildning.

Kilde: © EuroGeographics for de administrative grænser.

Endvidere gælder der generelt for regioner i dette område af Europa, at de har relativt mange naboer, da regionsstørrelserne er relativt små i forhold til andre områder.

Resultatet for Moran I-testen om rumlig afhængighed er præsenteret i Tabel 9.

Vægt	Moran I	M-I statistik	P-værdi
W_1	0,452	12,36	$2,2 \cdot 10^{-16}$
W_2	0,226	8,0192	$1,065 \cdot 10^{-15}$

Kilde: Egne beregninger.

Testresultatet viser, at p-værdien for Moran-testen med begge vægte ligger under signifikansniveauet på 5 pct., hvilket betyder at nullhypotesen om ingen rumlighed afhængighed afvises. Den rumlige autokorrelation er højere indenfor 205 km. distancen da Moran I er lig 0,452 for W_1 og er derved større end Moran I for W_2 som er lig 0,226. De positive fortegn på Moran I indikere at observationer i datasættet med lignende værdier har en tendens til at være geografisk koncentreret, dvs. at de optræder som naboer. Dette betyder at den lineære regressionsmodel ikke kan give unbiased parameterestimer, hvorfor der må anvendes andre modelspecifikationer, som kan håndtere den rumlige afhængighed. Derfor estimeres i følgende afsnit en SDM og en SDEM, som blev beskrevet i afsnit 4.3.1.

5.2 Estimationsresultater

Estimationsresultaterne for de forskellige modeller er præsenteret i Tabel 10. Model 1 er den lineære model, hvor der ikke er taget højde for rumlig afhængighed, mens model 2 og 3 er specificeret som en SDM, og de har hhv. vægtene W_1 og W_2 . Model 4 og 5 er specificeret som en SDEM, som hhv. har vægtene W_1 og W_2 . Parameterestimerne for model 2 og 3 kan, til forskel fra de andre modeller, ikke fortolkes direkte som beskrevet i afsnit 0, hvorfor de direkte og indirekte effekter er beregnet og præsenteret i hhv. Tabel 11 og Tabel 12. Parametrene for de tre første variable måler de direkte effekter, mens parametrene for de laggede variable måler de indirekte effekter.

Tabel 10 Estimationsresultater

	Model				
	1	2	3	4	5
	OLS	SDM, W_1	SDM, W_2	SDEM, W_1	SDEM, W_2
hrst02	0,002332 . (0,001194)	0,00083934 (0,0010444)	0,0011020 (0,0010494)	0,00080971 (0,0010205)	0,0019035 (0,0011692)
tPAT02	0,0002526*** (0,00006304)	0,000074265 (0,000054827)	0,00016903** (0,000055323)	0,00010803** (0,000052488)	0,00016556** (0,000059610)
dcommu	0,001908 (0,01528)	-0,023101 . (0,012110)	0,0026141 (0,013140)	-0,015303 (0,012042)	0,00074102 (0,014271)
Lag hrst02		-0,0010401 (0,0017270)	0,0024515 (0,0026869)	0,0029219 (0,0027865)	0,0052360 . (0,0030853)
Lag tPAT02		0,00020293 . (0,00011163)	-0,00013909 (0,00012307)	0,00040327** (0,00016808)	-0,00030367 (0,00015146)
Lag dcommu		0,070469** (0,023262)	-0,0058896 (0,035716)	0,13901*** (0,040130)	-0,011735 (0,042572)
rho		0,71167***	0,7475***		
Lambda				0,76684***	0,76507***
AIC		-417,03	-355,16	-422,87	-355,78
AIC lm		-339,81	-328,66	-349,81	-328,66

Signifikansniveau: 10 %, 5%, 1%, >1% for hhv. '.', '**', '***' og '****'

HRST02: human ressourcer i S&T, tPAT02: totale patenter, dcommu: udviklingen i antal pendlere, Proxy for hhv. tavs viden, kodificeret viden, samt kontrolvariablen dcommu.

Lm = lineær model

Model 1 viser den lineære regressionsmodel, hvor det ses at både tavs og kodificeret viden er signifikante, og har positive fortegn, hvorfor de begge har en positiv effekt på væksten i velstand. Pendling er derimod ikke signifikant, hvorfor der ingen effekt er på en regions vækst, hvis flere folk pendler ud af regionen for at arbejde. Parameterestimerne er dog biased idet der, jf. forrige afsnit, er rumlig autokorrelation mellem residualerne. Ved en umiddelbar inspektion af model 2, er de direkte og indirekte effekter af pendling signifikante, og ligeledes er de indirekte effekter af kodificeret viden. I model 3 er det kun de direkte effekter af kodificeret viden der er signifikante. Den endelige konklusion vil dog først kunne drages efter yderligere beregninger af de direkte og indirekte effekter. I model 4 ses det at både de direkte og indirekte effekter af kodificeret viden er signifikante sammen med de indirekte effekter af pendling. Dvs. regioners egen produktion af kodificeret viden har en positiv effekt på vækst, men ifølge modellen har de indirekte effekter af kodificeret viden en større effekt på væksten end de direkte, idet den har en større parameterværdi. De positive indirekte effekter betyder at der eksisterer positive spillovers af kodificeret viden mellem regioner indenfor en afstand af ca. 205 km. Som forventet eksisterer der en positiv indirekte pendlingseffekt, fordi der er en

sandsynlighed for at region i modtager nogle af region j 's pendlere, som derved bidrager positivt til væksten i region i i stedet for i deres bopælsregion j . I model 5 ses det at både den indirekte effekt af tavs viden og direkte effekt af kodificeret viden er signifikante. Modellen viser således at der er signifikante spillover af viden mellem region i og naboregionerne indenfor et bånd på 205 og 340 km, hvilket ikke er forventet.

Rho (ρ), som indgår i SDM specifikationerne, er et udtryk for den laggede værdi af den afhængige variabel, som dækker over, hvor stor en effekt andre regioners vækst har på region i 's vækst. Rho er positiv og signifikant i både model 2 og 3, dvs. at en regions vækst påvirkes positiv af en stigning i andre regioners vækst.

Lambda (λ), som indgår i SDEM specifikationen, er et udtryk for rumlig afhængige fejllid. Lambda er positiv og signifikant for både model 4 og 5, hvilket indikerer at de rumlige modeller er bedre end en lineær model, pga. den rumlige afhængighed.

Ud fra AIC som modelspecifikationskriterium ses det at SDM og SDEM passer bedre på data end en lineær model, idet at AIC er mindre hos model 2-5 end AIC for den lineære model.

For at kunne drage nogle egentlige konklusioner for de enkelte variables effekter i model 2 og 3 er de direkte og indirekte effekter beregnet og præsenteret i Tabel 11 og Tabel 12.

Tabel 11 direkte og indirekte effekter model 2 (W_1)		
Variable	Estimer	P-værdi
hrst02		
Direkte	0,0007660148	0,537990
Indirekte	-0,001462189	0,738277
Total	-0,0006961743	0,8726926
tPAT02		
Direkte	0,0001166296	0,031004
Indirekte	0,000844757	0,014176
Total	0,0009613866	0,0065883
dcommu		
Direkte	-0,0141521973	0,289408
Indirekte	0,178435022	0,019200
Total	0,1642828245	0,0388791

Kilde: Egne beregninger

Hverken den direkte eller indirekte effekt af tavs viden er signifikant, indenfor 205 km. båndet. De direkte og indirekte effekter af kodificeret viden er derimod positive og signifikante, da p-værdien er under signifikansniveauet på 5 pct. Dvs. at en stigning i kodificeret viden har en positiv totaleffekt på væksten og den totale effekt er ligeledes signifikant. Den direkte effekt af

udviklingen i antal pendlere er ikke signifikant, det er den indirekte derimod. Den indirekte effekt dominerer den direkte effekt, hvilket gør den totale effekt positiv og signifikant.

Tabel 12 Direkte og indirekte effekter model 3 (W_2)

W_2	estimerer	P-værdi
hrst02		
Direkte	0,0013914011	0,286228
Indirekte	0,01268172	0,26099
Total	0,0140731260	0,23580
tPAT02		
Direkte	0,0001679025	0,013076
Indirekte	-0,00004934158	0,84116
Total	0,0001185609	0,94582
dcommu		
Direkte	0,0022663565	0,832869
Indirekte	-0,1523846	0,96065
Total	-0,0129721080	0,97404

Kilde: Egne beregninger

Mellem region i og dens naboregioner indenfor båndet 205 til 340 km., er det kun den direkte effekt af kodificeret viden der er signifikant. Der er ingen af de indirekte effekter der er signifikante, hvilket indikerer at der ingen spillover-effekter er ved dette bånd. Model 3 og 5 er således uenige om, hvorvidt der er spillover af tavs viden ved dette bånd.

Opsummering

Den indledende test af rumlig afhængighed har vist, at der er en rumlig afhængighed mellem regionerne i EU-15 lande. Denne afhængighed er til stede mellem regioner, som befinder sig indenfor en distance på 205 km., og den er også til stede når region i 's naboregioner befinder sig mellem 205-340 km. væk. Afhængigheden falder dog med distancen, da den er større indenfor 205 km. På denne baggrund er der blevet estimeret to rumlige økonomiske modeller, som tager højde for førnævnte afhængighed. Estimationsresultaterne viser at kodificeret viden har en signifikant og positiv effekt på regional vækst. Der er ifølge resultaterne en tendens til, at den indirekte (spillover) effekt af kodificeret viden ikke er signifikant, når naboregionerne befinder sig i en afstand, der er større end 205 km. Modellerne opfanger overordnet ikke nogen signifikant effekt af tavs viden. Den indirekte effekt af udviklingen i antallet af pendlere, der pendler ud af bopælsregionen, er signifikant og positiv, og den lader til at opveje den direkte effekt som er negativ (og insignifikant), hvilket giver en positiv og signifikant total effekt.

5.2.1 Intra- og internationale spillover-effekter mellem regioner?

I ovenstående er der ikke taget højde for landegrænser, og derfor er det næste nærliggende spørgsmål, hvorvidt spillover-effekterne både eksisterer mellem regioner indenfor landegrænser, og mellem regioner på tværs af landegrænser. Derfor laves nu to nye vægtmatricer, som adskiller sig fra de forrige ved at de indeholder naboer hhv. indenfor og på tværs af landegrænser, og at der her kun laves en vægtmatrice for en distance på op til 205 km. Tabel 13 viser, hvor mange naboer det giver.

Tabel 13 Oversigt over naboskaber indenfor landegrænser og på tværs af landegrænser

Bånd	Antal naboskaber	Naboskaber i gennemsnit	[Min, Max]
<i>Indenfor landegrænse</i>	1700	10,24	[1,27]
<i>På tværs af landegrænser</i>	884	8,93	[1,33]

Note: Baseret på $n = 166$ regioner i indenfor landegrænse og $n = 99$ på tværs af landegrænser.

Kilde: Egne beregninger.

Med de to nye stikprøver foretages regressionerne nu for, at undersøge om landegrænser spiller en rolle for hvorvidt spillover-effekter stadigvæk er signifikante, når der ses på tværs af landegrænser. Resultaterne er angivet i Tabel 14 .

Tabel 14 Direkte og indirekte effekter for SDM (W_1)

Variable	Intranational		International	
	Estimerer	P-værdi	Estimerer	P-værdi
hrst02				
Direkte	0,0001181404	0,9896	0,0001050187	0,9203
Indirekte	0,0005635329	0,7933	-0,001862290	0,4105
Total	0,0006816733	0,8075	-0,001757271	0,3914
tPAT02				
Direkte	0,0000752806	0,1160	0,0002141566	0,0074
Indirekte	0,0006453441	0,0110	0,00002792727	0,9508
Total	0,0007206247	0,0068	0,0002420839	0,1844
dcommu				
Direkte	-0,009752086	0,3774	-0,0188395152	0,2945
Indirekte	0,1158237146	0,0320	0,04733996	0,2906
Total	0,1060716287	0,0680	0,0285004448	0,5702

Kilde: Egne beregninger

Resultaterne indenfor landegrænser viser, at der stadig er positive og signifikante spillover-effekter af kodificeret viden og af pendlere, mens tavs viden stadig er insignifikant. På tværs af landegrænser ser resultaterne meget anderledes ud. Her er der ingen signifikante spillover-effekter, og det er kun den direkte effekt af kodificeret viden der er signifikant. Konklusionen

for SDEM-modellen er den samme.¹⁷ Med disse resultater ses der en tendens til, at landegrænser fungerer som en barriere for interregionale spillovers effekt på vækst.

5.3 Robusthedstjek

For at undersøge resultaternes stabilitet er der udført en række robusthedstjeks, for at undersøge, hvor følsom modellen er over for modelspecifikationerne. I variabelbeskrivelsen blev der overvejet hvorvidt tertiær uddannelse eller antal ansatte i S&T skulle bruges som proxy for tavs viden. Anvendes tertiær uddannelse som en alternativ proxy for tavs viden, så ændrer det ikke på resultaterne. Tavs viden er stadig insignifikant. De andre variables signifikansniveau og fortegn afhænger heller ikke af, hvordan tavs viden er specificeret. Det er heller ikke afgørende for signifikansniveau eller fortegn, om tavs viden indgår i modellen eller ej.

Resultaterne er mindre robuste overfor valget af vækstperiode, dvs. afhængig af hvilken periode der vælges, varierer resultaterne i nogen grad, hvilket er en velkendt problematik i tværnsnitsdata. Vælges vækstperioden 2002-2012 fås de samme resultater som for vækstperioden 2002-2014. Vælges derimod vækstperioden 2002-2015 ændrer det meget på resultaternes fortegn og signifikans.

Yderligere kontrolleres der for udeladte variable, som i vækstlitteraturen anses for at være relevante for væksten. Der kontrolleres specifikt for graden af agglomeration, og hvorvidt der er tale om en hovedstadsregion eller ej. Ved at inkludere agglomeration (densi), så ændres resultaterne ikke for model 2 (SDM, W_1), idet fortegn og signifikansniveau er uændret for alle variable, men hverken den direkte, indirekte eller totale effekt af agglomeration er signifikant. I model 3 (SDM, W_2) ændres resultaterne ligeledes ikke ved at inkludere agglomeration, men til gengæld er den direkte effekt af agglomeration signifikant, men negativ. I modsætning til model 2 & 3, så ændres model 4 (SDEM, W_1) en smule af at inkludere agglomeration, idet estimatet for tavs viden nu bliver signifikant, men derudover er der ingen ændringer i fortegn eller signifikans for de andre variable.¹⁸ Det er ligesom før kun den direkte effekt af agglomeration, der er signifikant. Tilføjelsen af agglomeration i model 5 ændrer signifikansniveauet for tavs viden, idet den indirekte effekt før var signifikant, hvorimod det nu er den direkte effekt, der er signifikant.

¹⁷ For tabel over estimationsresultaterne for SDEM se bilag 2.

¹⁸ Signifikansen af tavs viden i model 4, er dog ikke robust overfor ændringer i vækstperioden.

Således ser det ud til at estimaterne for kodificeret viden ikke er følsomme overfor om agglomeration er inkluderet eller ej, men det er estimaterne for tavs viden i model 4 og 5 derimod. Dette sagt, falder AIC hos alle modellerne (2-5) ved at inkludere agglomeration, men det er ikke med mere end maksimalt seks enheder, hvilket vil sige at modellerne kun forbedres minimalt.

Top-10 listen for regioner med højeste niveau af hrst02 var jf. Tabel 5 stærkt domineret af hovedstadsregioner. Derfor er det relevant at kontrollere for betydningen af at være en hovedstadsregion. Dette gøres igennem en dummyvariabel, der tager værdien 1 hvis en region indeholder landets hovedstad, og 0 hvis regionen ikke indeholder en hovedstad. Ved at inkludere en hovedstadsdummy ændrer det ikke resultatet for model 2 og 3. Her er fortegnene de samme, og det ændrer heller ikke på hvilke variable der er signifikante. Hovedstadsdummyen selv er ikke signifikant i nogle af modellerne, og AIC er heller ikke lavere end modellen uden hovedstadsdummyen, hvorfor den ikke gør modellen bedre ved at inddrage den. For model 4 og 5 giver det stort set ingen ændringer. Den eneste ændring er, at den direkte effekt af tavs viden nu bliver signifikant i model 5. Igen er dummyvariablen ikke selv signifikant, og AIC for modellerne er heller ikke bedre, hvorfor variabelen ikke er medtaget i den endelige model.

Inkluderingen af hovedstadsdummi har således ikke nogen effekt på resultaterne for kodificeret viden. Det har en lille effekt for tavs viden, idet den direkte effekt bliver signifikant i model 5, men ellers ændrer det ikke ved noget.

5.3.1 Delstikprøve baseret på udviklingsniveau

Der er også blevet tjekket for hvor robuste estimaterne er, når regionernes forskellige udviklingsniveauer tages i betragtning. Regionerne i stikprøven er som sagt blevet opdelt i to mindre stikprøver, alt efter regionernes niveau for BNP per indbygger i vækstperiodens startår. Den første delstikprøve dækker de 25 pct. af observationerne i den fulde stikprøve, som ligger over tredje kvartil i BNP per indbygger, mens den anden dækker de 75 pct. af observationerne som ligger under tredje kvartil i den fulde stikprøve.

Top-25 stikprøve

Model 2 og 3 (SDM-modellerne)

Estimaterne for model 2 (SDM, W_1) har med top-25 stikprøven hovedsageligt fået de forventede fortegn og resultaterne er angivet i Tabel 15. Til forskel fra regressionen med den fulde

stikprøve, er det bemærkelsesværdigt at hverken den direkte eller indirekte effekt af kodificeret viden længere er signifikante indenfor 205 km.

Tabel 15 Top-25: direkte og indirekte effekter model 2 (W_1)

Variable	Estimater	P-værdi
hrst02		
Direkte	0,0001110541	0,9695
Indirekte	0,0024001840	0,5716
Total	0,002511238	0,5477
tPAT02		
Direkte	0,0001102816	0,2250
Indirekte	0,0005236213	0,1216
Total	0,000633903	0,0735
dcommu		
Direkte	0,0670586322	0,0247
Indirekte	-0,0215985513	0,8985
Total	0,045460081	0,7019

Kilde: Egne beregninger

Den direkte og totale effekt af kodificeret viden er dog signifikante for estimaterne i model 3 (SDM, W_2), hvilket ikke var tilfældet med den totale effekt under den fulde stikprøve. Ligeledes er den direkte og totale effekt af pendling signifikante og positive i model 3 modsat tilfældet med den fulde stikprøve, men det positive fortegn ved den direkte effekt var ikke ventet. Ifølge SDM-modellen har viden således i sin kodificeret og tavse form ikke nogen signifikant effekt på vækst indenfor 205 km., når der udelukkende ses på relativt mere udviklede regioner (top-25). Ydermere er $\rho = 0,33573$ for model 2, men den er ikke længere signifikant, dvs. nulhypotesen om at der ingen rumlig autokorrelation er i regionernes vækstrater, ikke kan afvises. Ikke overraskende afspejler dette sig i model 2's AIC på -68,03, idet denne værdi til sammenligning ikke er spor forskellig fra den lineære models AIC, som er lig -68,036. Dette betyder, at model 2 på baggrund af top-25 stikprøven ikke er bedre til at forklare data, end den lineære model til sammenligning er.¹⁹ Foretages en OLS regions på top-25 stikprøven er kodificeret viden og udviklingen i pendlere signifikant, men tavs viden er stadig ikke signifikant. Det samme gør sig gældende for model 3, hvor ρ ligeledes er insignifikant, og har en AIC lig -70,559, som er højere end AIC for den lineære model, som er lig -72,076, og OLS resultaterne giver ligeledes det sammen som for bånd 1.

¹⁹ Model 2 opnår naturligvis en højere AIC på top-25 stikprøven, end den gør for den fulde stikprøve, idet regressionen er baseret på en relativt mindre stikprøve. Derfor kan AIC ikke bruges til at sammenligne hvor god modellen er til at beskrive data på tværs af forskellige stikprøvestørrelser.

Model 4 og 5 (SDEM-modellerne)

Estimaterne for SDEM-modellerne differentierer sig også fra dem opnået med den fulde stikprøve, som angivet i Tabel 16.

Tabel 16 Top-25: estimationsresultater for SDEM

	4	5
	SDEM, W_1	SDEM, W_2
hrst02	0,00039945 (0,0025428)	0,0034526 (0,0024507)
tPAT02	0,00010035 (0,000079908)	0,00017330 * (0,000075863)
dcommu	0,069018 * (0,032390)	0,070028 * (0,029847)
Lag hrst02	0,0020349 (0,0038801)	-0,0014608 (0,0035726)
Lag tPAT02	0,00040941 * (0,00020292)	0,00021471 (0,00014776)
Lag dcommu	-0,0079493 (0,092735)	0,13464 . (0,080320)
Lambda	0,28299	-0,3071
AIC	-67,126	-70,834
AIC lm	-68,036	-72,072

Kilde: Egne beregninger

Estimaterne for model 4 (SDEM, W_1) har stort set opnået de forventede fortegn, men det er blot den indirekte effekt af kodificeret viden der er signifikant indenfor 205 km, og ligeledes er den direkte effekt af pendling signifikant, dog med positiv fortegn mod forventning. Model 4 har en AIC på -67,126, hvilket er højere end for den lineære model, som er lig -68,036, hvilket også indikerer at model 2 baseret på top-25 stikprøven ikke forklarer data bedre end den lineære model. Størrelsen λ er modsat regressionen på den fulde stikprøve heller ikke signifikant, hvorfor nulhypotesen om at fejlleddene ikke har en rumlig afhængighed, ikke kan afvises. Hos model 5 (SDEM, W_2) er det i stedet den direkte effekt af kodificeret viden, der er positiv og signifikant. Resultaterne for λ og AIC i model 5 er de samme som i model 4, dog med undtagelse af at λ nu er negativ.

Således er SDM og SDEM modellerne uenige om hvorvidt den direkte eller indirekte effekt af kodificeret viden er positiv for top-25 regioner, når deres naboregioner er defineret ud fra båndet på 205-340 km. Ydermere kan både SDM- og SDEM-modellerne ikke give en bedre forklaring af data i top-25, end den lineære model kan.

Bund-75 stikprøven

Model 2 og 3 (SDM-modellerne)

Der tegner sig et andet billede, når regressionen udføres på stikprøven med regionerne der har et relativt lavere udviklingsniveau (bund-75). Ifølge model 2 (SDM, W_1) har kodificeret viden nu en positiv og signifikant direkte, indirekte og total effekt på vækst indenfor 205 km. P-værdierne er på hhv. 0,0045, 0,06 og 0,026, dvs. resultaterne er signifikante indenfor hhv. 1 pct., 10 pct. og 5 pct. signifikansniveau. I samme forbindelse er den direkte effekt af pendling både signifikant og negativ, dvs. det har en negativ effekt på regional vækst hvis udviklingen i antal folk der pendler ud af regionen er stigende, hvorimod den indirekte (spillover) effekt både er positiv og signifikant, dvs. det har en positiv effekt på vækst, hvis der sker en positiv udvikling i antal folk, der pendler ud af naboregionerne. Der gælder desuden, at ρ er signifikant, og at model 2 med en AIC lig -352,25 bedre kan forklare data i bund-75 end den lineære model kan med en AIC lig -262,38 kan. Ifølge model 3 (SDM, W_1) er det blot den direkte effekt af kodificeret viden, der er signifikant blandt de forklarende variable, og ρ er ligeledes signifikant, og model 3 kan med en AIC lig -287,53 bedre forklare data end en lineær model kan med en AIC som er lig -254,25.

Model 4 og 5 (SDEM-modellerne)

Parameterestimerne i model 4 (SDEM, W_1) får de forventede fortegn i regressionen, når bund-75 udgør stikprøven. Både den direkte og den indirekte effekt af kodificeret viden er signifikant indenfor 205 kilometers distancen indenfor et signifikansniveau på hhv. 1 pct. og 10 pct. Den direkte og indirekte effekt af udviklingen i pendling er ligeledes signifikant indenfor et signifikansniveau på hhv. 5 pct. og 1 pct. Ydermere er $\lambda = 0,82670$ og signifikant, hvilket betyder at nulhypotesen om, at fejlleddene ikke har en rumlig afhængighed afvises. Modellen opnår en AIC-værdi på -357,08, hvilket til sammenligning er lavere og dermed bedre end AIC for den lineære model, hvor værdien er lig -262,38. I model 5 (SDEM, W_2) er den direkte effekt af kodificeret viden signifikant, og ligeledes er spillovers af kodificeret viden fra naboregionerne i båndet 205-340 km. signifikant, og det samme er den direkte og indirekte effekt af udviklingen i antallet af pendlere, som pendler ud af bopælsregionen. Ifølge model 5 er λ stadigvæk signifikant, og er bedre til at forklare data end den lineære model ville kunne.

Opsummering

Kodificeret viden har ifølge alle modellerne en positiv og signifikant direkte effekt på væksten i perioden 2002-2014 i de regioner, som ligger i bund-75 i relation til udviklingsniveauet i 2002. Således afviger konklusion for bund-75 ikke fra konklusionen for hele stikprøven. Resultatet er ikke lige så entydigt for top-25 regionerne, idet kodificeret viden i dette tilfælde kun lader til at være signifikant, når naboregionerne er defineret til at befinde sig i 205-340 km.-båndet jf. model 3 og 5. Det bemærkelsesværdige resultat er, at den rumlige afhængighed ikke er tilstede i top-25 regionerne. Udviklingsniveau lader således til at have en betydning for, om den direkte effekt af kodificeret viden er signifikant eller ej, men udviklingsniveauet påvirker til gengæld ikke resultaterne for tavs viden.

6 Diskussion

I nærværende afsnit vil estimationsresultaterne diskuteres, og efterfølgende vil resultaterne relateres til videnspolicy i EU.

6.1 Fortolkning og diskussion af resultater

Resultatafsnittet viste at viden og spillover af viden kan være med til at forklare nogle tendenser indenfor regional vækst. Ifølge de to estimerede modeller har kodificeret viden som forventet en positiv og signifikant direkte og indirekte effekt på væksten indenfor 205 km. Dette indikerer, at en regions vækst afhænger af egen vidensproduktion, men hvad mere interessant er, at den også er påvirket af naboregionenes vidensproduktion. At den direkte effekt er signifikant, er med til at understøtte Romers teori om, at udviklingen af viden har en effekt på væksten. En region kan derved opnå en stigning i velstanden ved at satse på F&U, for at fremme produktionen af kodificeret viden. Derudover indikerer den indirekte effekt, at kodificeret viden er i stand til at rejse på tværs af regionsgrænser, og derved kan naboregionerne udnytte den viden, der nu er tilgængelig. Samtidig viste Figur 9, at der er en umiddelbar tendens til at regioner med stor vidensproduktion også ligger geografisk tæt på hinanden. Dette kan tolkes som, at viden i disse områder rejser lettere mellem regioner, fordi regionerne må antages at have evnen til at absorbere den nye viden, når de også selv har en stor produktion af viden. Modellerne viser dog også, at der er en grænse for hvor langt væksteffekterne kan rejse, da spillover-effekten ikke længere er signifikant, når naboregionerne er længere væk. Når regionerne ligger længere væk fra hinanden, så er det kun den direkte effekt af kodificeret viden der er signifikant, og har en positiv på væksten, hvilket kan indikere at en region her er mere afhængig af sin egen produktion af kodificeret viden, når den ses i forhold til naboregioner der ligger langt væk, da de ikke opnår nogen spillover-effekt på væksten fra disse regioner. Dette er dog ikke et udtryk for, at kodificeret viden ikke kan rejse længere end de ca. 205 km, som udgør cut-off grænsen, men i stedet et udtryk for at der findes en grænse for, hvor langt væk spillover-effekter virker på væksten i andre regioner. Rodríguez-Pose & Crescenzi (2008) finder også, at der blandt de europæiske regioner eksisterer en geografisk grænse for, hvor langt viden kan rejse. De knytter det fund op på den overvejelse at transport af viden i høj grad er afhængig af menneskelig kontakt, og derfor ikke kan rejse længere end, hvad et menneske kan nå på en dag. Viden kan dog også spredes på andre måder. Viden, og især den kodificeret viden, har gode betingelser for at blive spredt i kraft af internettet og det virtuelle rum, men resultaterne her viser, at der stadig er en geografisk begrænsning for, hvor langt væk spillover-effekter er i stand til at bevæge sig,

hvorfor det stadig er relevant at betragte geografiske distance som en mulig barriere for videns spillover-effekter.

Tavs viden er i modellen ikke signifikant, og dette blev yderligere bekræftet af robusthedstjekkene. At den indirekte (spillover) effekt af tavs viden ikke er signifikant, er dog ikke helt mod forventning. Tavs viden er som beskrevet tilknyttet en person i form af uddannelse og erhvervserfaring, dvs. at hvis tavs viden skal rejse på tværs af regionsgrænser og derved bidrage til væksten, så kræver det at personen der besidder den tavse viden også bevæger sig på tværs af regionsgrænser. En regions vækst forventes i stedet i langt højere grad at være afhængig af egen beholdning af tavs viden, som manifesteret i mængden af humankapital, og netop den tavse viden iboende i humankapital udgør et vigtigt element der ifølge Romer er forudsætningen for overhovedet at kunne producere ny viden. Derfor er det mod forventning, at den direkte effekt ikke er signifikant for tavs viden. I et studie af Capello & Fratesi (2013) bruges antal ansatte i S&T i pct. af arbejdsstyrken også. De finder at variabelen er signifikant i modellen, som er estimeret på baggrund af alle regioner i EU-27. De får dog et andet resultat når de deler stikprøven op i to mindre stikprøver, hvor den ene består af EU-15 og den anden af de 12 lande som kom med ved østudvidelsen. De får nemlig at variabelen ikke længere er signifikant for EU-15, hvilket er i tråd med dette projekts resultater. Dette resultat kan skyldes flere forhold, men et afgørende forhold er, at tavs viden er svært at kvantificere og derfor svært at opfange. Endnu et muligt forhold er, at der arbejdes med NUTS2 regioner som administrative enheder. Jo mere disaggregerede niveau der arbejdes med, jo større sandsynlighed er der for at tavs viden kan opfanges. Således kunne der alternativt arbejdes med NUTS3 niveau eller med et endnu lavere niveau helt ned på virksomhedsniveau. Lavere niveauer forudsætter dog at der arbejdes med andre variable, da variable som antal ansatte i S&T eller tertiær uddannelse, som også er en meget brugt proxy for tavs viden, ikke findes på så lavt et niveau. De variable som på det niveau vil kunne være interessante i forhold til at beskrive tavs viden, såsom ekspertviden og effektivitet (Ryan & O'Connor, n.d, s. 6), ville dog ikke være tilgængelige for projektets forfattere. Det administrative niveau har betydning for i hvor stor grad spillover-effekten af tavs viden kommer til udtryk. Da det er gennemsnittet af hele regionens værdier, der tages udgangspunkt i, har det en betydning om der arbejdes på lande niveau eller NUTS2 niveau, fordi variationen vil være større på højere niveau, hvormed mere information går tabt. Det kunne dog tænkes at byer langs regionsgrænsen havde en meget større påvirkning af byer på den anden side end hvad regionen som helhed har. Derfor giver det mening at disaggregere så meget så muligt,

hvorfor NUTS2 er valgt som administrativ enhed. Dette gør at projektet kan gå dybere, men stadig holde sig på et makroniveau. En anden afgørende ting er, at dette studie undersøger videns effekt på vækst, men som Tabel 5 og Tabel 7 viste, var mange af de regioner som har en vidensintensiv arbejdsstyrke også de regioner, der havde et højt udviklingsniveau. Regioner der i forvejen har et højt udviklingsniveau har ikke haft lige så store vækstrater, men når det samtidig er de regioner, der har den største andel af arbejdsstyrken der arbejder i S&T, er der risiko for at effekten af tavs viden ikke opfanges, fordi tavs viden på den måde ikke forbindes med høj vækst. Af disse regioner udgør hovedstadsregioner en stor andel. For at undersøge om det forholder sig sådan, er Tabel 17 konstrueret. Her er hovedstadsregionernes placering indenfor variablene angivet, og de er rangeret fra 1-167, hvor 1 er den bedst placeret region, mens den dårligste placeret region har nr. 167.

Tabel 17 Hovedstadsregionernes placeringer indenfor variablene.

Hovedstadsregioner	Vækst	Placering		
		Ansatte i S&T	patenter	udviklingsniveau
Vienna (AT)	107	42	35	6
Bruxelles-Cap (BE)	155	5	70	1
Ile-de-France (FR)	51	4	17	5
Berlin (DE)	19	7	40	69
Southern and Eastern (IR)	61	74	116	14
Lazio (IT)	164	90	140	23
Luxembourg (LU)	12	81	52	2
Noord-Holland (NL)	85	6	64	9
Lisbon (PT)	73	154	161	67
Madrid (ES)	81	9	145	31
Stockholm (SE)	48	1	15	11
London (UK)	47	8	101	4
Eastern Scotland (UK)	136	46	84	56
East Wales (UK)	138	85	112	87

Kilde: Egne beregninger

Tabellen viser en tendens til, at regioner med relativt høje udviklingsniveauer placerer sig relativt dårligere i forhold til vækst. Luxembourg har dog som den eneste formået at have et højt udviklingsniveau, og samtidig have en relativ god placering i forhold til vækst væksten.

Denne observation kan også være med til at forklare, hvorfor det ikke bidrager til modellen at kontrollere for både hovedstadsregioner og befolkningstætheden. I Tabel 5 og Tabel 7 var der ellers en tendens til at hovedstadsregioner klarede sig godt indenfor disse variable, hvilket bl.a. var baggrunden for at kontrollere for det. Igen har disse regioner haft en relativ mindre vækst, som det fremgik af Tabel 17, hvorfor det kommer til at fremstå som om det påvirker væksten negativt at være en hovedstadsregion, selvom vi ser at de klare sig godt. Hovedstadsregioner

indeholder oftest også mange mennesker på relativt lidt plads. At befolkningstætheden ikke er signifikant, kan bl.a. skyldes førnævnte dilemma. Det kan også skyldes at de positive og de negative eksternaliteter ved at være mange mennesker samlet på et sted, udlignes af hinanden, hvorfor modellen ikke fanger nogen overordnede effekt.

Det har ligeledes også en betydning for væksten at udviklingen i pendlere stiger for nabo regionen. Intuitionen bag dette er, som beskrevet i variabelbeskrivelsen, at hvis folk pendler ud af naboregionen, er der sandsynlighed for at en andel af dem vil pendle ind i region *i*. Derfor er der som forventet en positiv indirekte effekt af udviklingen i pendlere. Den direkte effekt er derimod ikke signifikant, dog negativ som forventet. Den totale effekt er endvidere også positiv og signifikant. Dette indikerer at den positive indirekte effekt er eventuel større end den direkte, hvilket betyder at der er en overordnet positiv effekt af udviklingen i pendlere. Pendling er ikke kun interessant med henblik på at kontrollere for at en persons værdiskabelse ikke altid lægges i den region, hvor personen bor. For når personer rejser på tværs af regionsgrænser for at arbejde, så tager personen også sin tavse viden med sig. Det har altså, på baggrund af antagelsen om at en andel af andre regioners pendlere vil pendle ind i regionen, ikke overraskende, en positiv effekt på en regions vækst, hvis folk fra naboregionen rejser på tværs af regionsgrænser og bruger deres tavse viden i regionen, og den er større end den tavse viden der går tabt ved at regionens egne indbyggere pendler ud.

Valget af en Spatial Durbin model, forudsætter at der arbejdes med en række antagelser, som har implikationer for fortolkningen af resultaterne. I relation til dette projekt er der især én der giver modellen nogle begrænsninger. I beskrivelsen af de afledte effekter ligger da implicit den antagelse, at en ændring i de uafhængige variable får en øjeblikkelig effekt på væksten. Det betyder også at de indirekte effekter, altså spillover-effekterne, sker simultant med de direkte effekter (LeSage & Pace, 2009, s. 195). Intuitivt vil det dog ikke gøre sig gældende i virkeligheden, da det må antages at det vil tage tid inden ændringer i de uafhængige variable vil have en effekt på den regionale vækst. I dette studie gør dette sig især gældende for kodificeret viden udtrykt i form af patenter. Der kan som beskrevet i variabelbeskrivelsen gå mange år inden patenter slår igennem på væksten, da det produkt de har patenteret først skal produceres, kommerialiseres og til sidst skal forbrugerne tage produktet til sig. Ligeledes gælder det for virksomheder, at når de har patenteret en proces, så tager det tid inden den nye produktionsproces er implementeret. Modellen derimod antager at de direkte effekt sker med det

samme. I denne sammenhæng afhænger det også af hvilke variable der er inddraget. Med patenter er det en stærk antagelse fordi det vides at det tager tid før det har nogen målbar effekt på væksten, der er dog andre uafhængige variable, hvor antagelsen ikke er lige så stærk. Eksempelvis ville ændringer i beskæftigelsen relativt hurtigt kunne give en effekt på væksten. Antagelsen omkring tidsaspektet i forhold til de direkte effekter betyder også at feedback effekterne antages at ske samtidig med de direkte effekter. I afsnit 4.3.2 blev det beskrevet hvordan en ændring i region i har en effekt på region j 's vækst, som så igen vil påvirke region i 's vækst. Denne feedback proces må også forventes at tage tid, og især når det er viden der skal krydse regionsgrænsen, om det så er i kodificeret eller tavs form. For selvom kodificeret viden kan spredes relativt hurtigere kræver det stadig tid, inden naboregionen har tilegnet sig kvalifikationer til at kunne absorbere denne nye viden. Samme argument kan bruges i henhold til tidsaspektet for de indirekte effekter. Studiets anvendte metode har således sine begrænsninger, og disse kan især tilskrives at der tages udgangspunkt i en tværsnitsmodel. Egenskaben ved tværsnitsdata er, at den ikke inkluderer en tidsdimension, men i stedet opererer med et udsnit af en bestemt tidsperiode, hvorfor resultater også skal tolkes efter dette, og derved ikke tilskrive dem større betydning end, hvad de har. Der gives altså blot et øjebliksbillede af kodificeret- og tavs videns effekt på regional vækst.

6.2 Overnational videns- og vækstpolicy

Trods åbne landegrænser i EU, så er den geografiske rækkevidde for spillover-effekter af viden både begrænset i en kilometermæssig og i en landemæssig forstand. Regioners lokale bestræbelser, kombineret med interregionale interaktioner, ser ikke ud til at have udlignet de europæiske regioners udviklingsniveauer. Deraf er der en ujævn fordeling af højvækst- og lavvækstregioner. I teorien burde de mindre vidensintensive regioner kunne nærme sig de mere udviklede regioner via vækst, hvis de øger deres produktion af viden, idet der lader til at være en positiv sammenhæng mellem vidensproduktion og vækst, og ikke mindst udviklingsniveau. I teorien burde de også kunne tilnærme sig de mere vidensintensive regioner, altså med Boschma's (2005) ord øge den kognitive nærhed, ved at satse mere på viden, for at kunne udnytte interregionale spillovers af viden. I praksis har regionerne i Europa dog en udfordring ved, at de tilhører lande med forskellige sprog, kulturer og normer, idet viden kun kan bruges hvis den kan fremskaffes, forstås og kommunikeres. I dette henseende er *eksistensen* af viden ikke ensbetydende med *udnyttelse* af viden, hvad implikationen er, af endogen vækstteori jf. Cowan et al. (2000). Således er udfordringerne ikke kun intraregionale, men de er også

interregionale, da vidensudveksling involverer flere parter. På et overnationalt niveau er udfordringen således, at skabe en form for samhørighed mellem regioner, som i essensen er forskellige, men det kræver også støtte på et nationalt niveau. EU har på et overnationalt niveau forsøgt at imødekomme disse udfordringer, men givetvis med medlemslandenes hjælp.

Overordnet har EU en regionalpolitik, der sigter på at fremme samhørigheden mellem de europæiske regioner ved at udligne forskellene mellem de rige og de fattige regioner. For at kunne gennemføre denne regionalpolitik, så har EU afsat 2,607 mia. kr. i perioden 2014-2020, og ydermere påkræver EU, at landene er med til at selvfinansiere en del af politikken, således at budgettet forventes at komme på 3,352 mia. kr. (EU-oplysningen, 2017). EU's regionalpolitik lægger beslag på ca. en tredjedel af EU's samlede budget (EU-oplysningen, 2017), og fylder således meget hos EU rent økonomisk og policymæssigt. EU's finansiering sker gennem fire såkaldte Strukturfonde²⁰ og en Samhørighedsfond. Gennem disse fonde gives der støtte til infrastrukturprojekter inden for eksempelvis transport og kommunikation, og der gives støtte til at små og mellemstore virksomheder kan blive mere innovative og konkurrencedygtige (EU-oplysningen, 2017).

EU har med deres regionalpolitik især prioriteret at skabe vidensbaseret jobs og vækst, som ikke kun skal tilgodese ledende forskningsregioner, men også mindre udviklede områder i Europa (Borch, et al., 2015, s. 20). Et initiativ har været at skabe ”*The European Research Area*” (ERA), altså et europæisk forskningsområde, som sigter på at skabe et indre marked for viden, forskning og innovation, for at gøre forskere, institutioner og virksomheder i stand til at cirkulere, konkurrere og samarbejde på tværs af grænser (Borch, et al., 2015, s. 20). Dette skulle i samarbejde med investeringer i diverse projekter øge mobiliteten af arbejdere, og give incitament til øget international samarbejde af aktører på tværs af europæiske regioner. Således lader det til, at Romers (1990) idé om en beholdning af viden, er forsøgt at blive operationaliseret i praksis med ERA, dog i en større international skala. Dette er i tråd med Cowan et al. (2000)'s konstatering, at eksistensen af en såkaldt vidensbeholdning forudsætter at viden bliver delt mellem økonomiske aktører. Det nærliggende spørgsmål er dog, hvor villige aktører er til at dele viden eller skabe viden sammen?

²⁰ Herunder Den Europæiske Udviklings- og Garantifond for Landbruget, Den Europæiske Fond for Regionaludvikling, Den Europæiske Socialfond og Fiskerifonden (EU-oplysningen, 2017).

På den ene side observeres der visse trends, som lader til at indikere, at vidensproducenter er samarbejdsvillige. Ifølge Borch, et al. (2015, s. 15) er der indikationer på, at der er en stigende afhængighed mellem forskere på tværs af landegrænser, idet 56 pct. af forskere i European Research, som er en samling af 28 nationale forskningssystemer, har brugt mindst tre måneder udenfor deres landegrænser på grund af forskningsrelateret arbejde. Cirka en femtedel af alle videnskabelige artikler bliver desuden medfattet internationalt, og for forskere i Frankrig, Tyskland og Storbritannien er det omkring halvdelen af deres artikler dette gør sig gældende for (Borch, et al., 2015, s. 25).

På den anden side er det også tydeligt, at visse aktører ikke ønsker at dele ud af sin viden og lade andre bruge den. Tidligere nævntes, at patenter bliver brugt til at afholde potentielle konkurrenter fra at bruge ens viden, og tage ens markedsandele. Dette kan på den anden side ikke demotivere forsøg på at udbrede co-produktion af viden og vidensdeling. Der eksisterer hindringer for, at viden bliver udbredt gnidningsfrit af diverse årsager, men disse kan i stedet forsøges at blive fjernet. Borch, et al. (2015, s. 7) anbefaler eksempelvis, at regler vedrørende intellektuel ejendomsret reformeres. Intellektuel ejendomsret har aldrig nogensinde været så forsvaret, som de er nu til dags med IKT-virksomheder, som tager hinanden til retten oftere og oftere, fordi deres produkter ligner hinanden (Borch, et al., 2015, s. 18). Der sker således en form for ressourcspild, når virksomheder bruger penge på at trække hinanden i retten, i stedet for at bruge dem produktivt i et andet henseende. En virksomhed skal på den ene side kunne nyde gevinsten af at investere store summer penge i viden, men på den anden side spørger Borch, et al. (2015, s. 52) : *hvor meget skal en virksomhed eksempelvis kunne opkræve for sin livsredende medicin, og hvor meget kan virksomheden kæmpe for at have den eksklusive ret til brugen af den viden, der er manifesteret i denne medicin?* Det er et godt spørgsmål, da produktionen af viden kræver store summer af penge, og hvis viden kun bliver produceret ud fra et profitmotiv, så skal vidensproducenter kunne kompenseres økonomisk for deres arbejde, ellers vil der potentielt ikke genereres viden. Ligesom der blev påpeget af Romer (1990), så må økonomiske agenter have incitament til at investere deres tid og penge i at skabe viden. Der må dog være en øvre grænse, da en økonomi i visse henseender kan have større gavn af at lade nogle former for viden være offentlige goder. Et godt eksempel er frit udbredelse af viden gennem uddannelse. En veluddannet arbejdsstyrke skaber grobund for generering af viden, men hvis uddannelse er forbeholdt en lille skare af folk, så ville en økonomi ikke kunne realisere sit potentiale og generere mere viden og velstand.

Således vil EU gerne hjælpe tilbagestående regioner med at øge velstanden, og EU identificerer også viden og vidensdeling som en vigtig faktor, og prøver at gøre forholdene for vidensdeling blandt de europæiske regioner lettere ved at oprette ERA, men al viden antager ikke form af et offentligt gode, hvorfor der stadig foreligger plads til yderligere tiltag.

7 Konklusion

Økonomer og politikere har længe været optaget af at skabe vækst for at øge velstanden. Der er dog en geografisk ujævn fordeling af hvor væksten skabes, og hvor udviklingsniveauet er højest blandt de europæiske regioner. På baggrund af moderne økonomiers stigende afhængighed af viden, og de observerede geografiske forskelle i vidensproduktionen, så er følgende problemformulering blevet dannet:

”Hvilken effekt har viden og interregionale vidensspillovers på regioners vækstperformance?”

Denne problemformulering er besvaret kvantitativt. Der er indsamlet data for regionerne i EU-15 lande, og disse er blevet anvendt til at estimere to såkaldte *rumlige* økonometriske modeller, som er blevet opstillet på baggrund af indsigter fra Romers (1990) teori om endogen vækst, regional udviklingslitteratur og tidligere forskning. Modellernes specifikationer er hhv. kendt som en *Spatial Durbin Model* og en *Spatial Durbin Error Model*. Det særlige ved disse modeller er, at de inkorporerer den geografiske dimension, og kan modellere interregionale spillover-effekter, hvilket åbner op for, at modellere regional vækst, således at væksten kan påvirkes af regionernes egen produktion af viden, såvel som andre regioners produktion af viden.

Implikationen af dette er, at regioner også er afhængige af de omkringliggende regioner, og derfor ikke skal ses i isolation, som der påpeges i regional udviklingslitteratur. Forskellen mellem modellerne er, at SDM muliggør at tage højde for andre regioners vækstrater, mens SDEM muliggør at tage højde for andre regioners fejlled. Idet viden er et bredt begreb, så er der i dette speciale skelnet mellem to typer af viden, herunder kodificeret viden og tavs viden, fordi de opfanger to vigtige aspekter af viden. Tavs viden er en form for viden, der har svært ved at blive spredt mellem mennesker, da den bygger på tillærte færdigheder, der er opnået gennem erfaringer, mens kodificeret viden har lettere ved at blive spredt, da det er en eksplicit form for viden. Implikationen af denne skelnen er, at den geografiske rækkevidde for interregionale spillover-effekter er begrænset, da diffusion af viden afhænger af i hvilken udtrækning viden kan kommunikeres, og i hvilken udstrækning bærerne af denne viden kan flytte sig. Derfor er udviklingen i antal folk, der pendler ud af en region også inkluderet, med henblik på at finde ud af hvor meget det gavner eller hindrer en regions vækst, at folk ikke arbejder, og derved udnytter den viden de har, i den region de bor i.

Ved hjælp af den økonometriske metode, er effekten af viden og interregionale viden spillovers på vækst således blevet estimeret for de europæiske regioner, og den geografiske rækkevidde af denne effekt er i denne sammenhæng blevet undersøgt. Estimationsresultaterne for både SDM og SDEM viser, at en regions lokal produktion af kodificeret viden har en positiv og signifikant effekt på vækst, og at interregionale spillovers af kodificeret viden har en positiv og signifikant effekt på vækst indenfor en rækkevidde på 205 km. Effekterne er dog ikke længere signifikante for en rækkevidde over 205 km. Resultaterne viser ydermere, at spillovers af kodificeret viden indenfor 205 km. ikke har en signifikant effekt på væksten i regioner på tværs af landegrænser, men at den har en signifikant effekt på væksten i regionerne indenfor landegrænserne. Begge modeller finder desuden, at den indirekte effekt af udviklingen i antallet af folk, der pendler ud af deres bopælsregion, har en positiv effekt på vækst. Det vil sige, at det har en positiv effekt på en regions i 's vækst, hvis der er flere der pendler ud af dens naboregioner, da en større andel af dem sandsynligvis ender med at arbejde og skabe værdi i region i . Estimationsresultaterne for tavs viden er overordnet ens for SDM og SDEM. Resultaterne viser imod forventning, at det lokale niveau af tavs viden ikke har nogen signifikant effekt på vækst indenfor en rækkevidde på 205 km., men som forventet er den insignifikant udover en rækkevidde på 205 km. Resultaterne divergerer for spillover-effekten af tavs viden. I SDEM er spillover-effekten mod forventning signifikant mellem en region i og dens naboregioner, som ligger indenfor et bånd, der er ca. 205-340 km. væk fra region i .

Der eksisterer således en positiv sammenhæng mellem viden, i sin kodificerede form, og vækst. Det er dog ligeledes tydeligt, at regioner der har en relativ høj produktion af viden, ikke nødvendigvis oplever de højeste vækstrater, men til gengæld fremgår det, at de i stedet har nogle af de højeste udviklingsniveauer. Dette fremgik både af den visuelle inspektion af det deskriptive data i afsnit 4.5.1.1 og 4.5.2.1, og senere af afsnit 5.3.1, hvor estimaternes robusthed overfor regionernes forskellige udviklingsniveau blev tjekket. Ligeledes har spillover-effekter af viden en positiv effekt på vækst, hvilket demonstrerer en gensidig afhængighed regionerne imellem.

8 Perspektivering

Modellernes specifikationer er robusthedstjekket, for at se om små ændringer ville skabe afgørende ændringer i resultaterne og derved den endelige konklusion. Ændringerne har alle været relativt simple, eksempelvis blev der tjekket om resultaterne var robuste overfor valg af tidsperiode. To helt grundlæggende specifikationer er der derimod ikke tjekket for. Disse to specifikationer er definitionen af naboer, altså konstruktionen af vægtmatricen, og niveauet af de administrative enheder der udgør observationerne. Valget af specifikationerne er truffet på baggrund af en lang række overvejelser. En ændring i disse to specifikationer, kan potentiel rykke ved resultaterne.

Et af de første valg der foretages, når der arbejdes med rumlig økonometri er, hvordan naboer skal defineres. Naboer skal ikke nødvendigvis forstås bogstaveligt, idet en nabo ikke behøver grænse op til en selv, som ellers er den normale fortolkning af ordet. Det hele afhænger af hvordan nærhed defineres. Nærhedsprincippet er et udtryk for, hvor tæt to regioner er på hinanden, og her er det ikke kun i en geografisk forstand sådan som dette projekt definerede det, men kan også være udviklingsmæssig eller teknologisk forstander. I dette projekt er der taget udgangspunkt i geografisk nærhed, altså hvor langt der, rent km. mæssigt, er mellem to regioner. Der er dog også andre former for nærhed, som f.eks. institutionel nærhed, social nærhed og kognitiv nærhed som blev nævnt tidligere (Boschma, 2005, s. 62). Det mest relevante nærhedsprincip i forhold til dette projekt er teknologisk nærhed. Med teknologisk nærhed, forstås hvor langt regionernes teknologiniveau ligger fra hinanden. Viden har bedre betingelser for at sprede sig, hvis regionerne, der skal udnytte denne nye viden, har teknologiske ligheder med den region, hvor den nye viden er opstået. Dette skyldes at hver form for teknologi har sit eget tekniske sprog og kræver forskellige egenskaber. Derfor må spillover af viden forventes at være højere mellem regioner med sammenlignelige teknologiske profiler (Greunz, 2003, s. 3). Dette kunne f.eks. gøre sig gældende for bilindustrien i Tyskland, der holder til i forskellige regioner, som ikke alle nødvendigvis er geografiske naboer, men det betyder ikke at viden ikke spredes på tværs af disse regioner. En meget populær²¹ måde at definere teknologiske naboer på, er ved at tage hver regions antal ansøgte patenter og dele dem op i forskellige klassifikationer, og finde hvor stor en del af de samlede patentansøgninger de udgør. Herefter sammenlignes hver region ud for hver patentklasse, og de rangeres på en skala fra 0 og 1, og jo mere regionerne

²¹ Der findes andre måder at definere teknologisk distance på. Denne måde stammer fra Jaffe (1986) som anses for at være en af de første til at beskrive teknologisk distance (Greunz, 2003, s. 3).

ligner hinanden teknologisk, jo tættere på 1 er de. Herefter udgør de tre regioner, som har de højeste værdier 1. ordens regioner, de tre næste regioner udgør 2. ordens regioner osv. Her er det vigtigt at påpege at naboer her ikke skal forstås som to regioner der også er geografiske naboer, på denne måde kan en engelsk og en tysk region også være naboer (Greunz, 2003, s. 7).

Teknologisk distance ville også kunne måles efter, hvor vidensintensiv arbejdsmarkedet er, f.eks. ved at bruge antal ansatte i S&T i pct. af arbejdsstyrken. Regionerne ville herefter blive inddelt i intervaller, og regionerne i samme interval ville være hinandens naboer. Havde dette projekt i stedet brugt teknologisk distance i stedet for geografisk distance, havde det åbnet op for muligheden for at vidensintensive regioner kunne være naboer uanset geografisk afstand. Ud fra europakortene fremgik det at hovedstadsregionerne generelt lå i top ved de fleste variable anvendt i projektet. Havde den teknologiske distance været brugt, ville nogle af disse regioner med stor sandsynlighed være 1. eller 2. ordens naboer, især hvis antal ansatte i S&T i pct. af arbejdsstyrken var blevet brugt til at definere naboer. Om det havde givet et andet resultat er dog ikke til at sige. For selvom teknologisk distance måske heller ikke ville have givet et signifikant resultat af tavs viden, er det mere et udtryk for at tavs viden for det første kan være svært at opfange, idet det også består af færdigheder der ikke bare kan overføres til andre, og for det andet kan det være svært at finde den rette proxy for tavs viden. Om teknologisk distance kan erstatte geografisk distance er der stor uenighed om. Som beskrevet i afsnit 2.2.3, mener nogle at geografisk distance ikke længere betyder noget, og at andre nærhedsprincipper som f.eks. teknologisk distance bør anvendes supplerende. En af de store forskelle ved disse to nærhedsprincipper er, hvilke regioner der kommer til at udgør naboer. En region kan ikke selv vælge sine geografiske naboer, de er placeret hvor de nu engang er placeret. Med teknologisk distance er regionen ikke længere afhængig af hvem den er omringet af, her vil regionen i stedet få naboer der har samme teknologiske profil som regionen selv. Samtidig er der den mulighed, at regioner kan arbejde hen imod en bestemt profil, hvis de ønsker det, og der er potentiale for det.

9 Litteraturliste

- Anselin, L. (2010). Thirty Years of Spatial Econometrics. *Papers in Regional Science*, 89(1), s. 3-25.
- Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3), s. 155-173.
- Audretsch, D. B., & Feldman, M. P. (2004). Knowledge Spillovers and the Geography of Innovation. I J. V. Henderson, & J. F. Thisse, *Handbook of Urban and Regional Economics* (Årg. 4, s. 2713-2739). Amsterdam: Elsevier.
- Balland, P.-A., & Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, 93(1), s. 1-23.
- Basile, R., Capello, R., & Caragliu, A. (2011). Interregional Knowledge Spillovers and Economic Growth: The Role of Relational Proximity. I R. Basile, K. Kourtit, P. Nijkamp, & R. R. Stough, *Drivers of Innovation, Entrepreneurship and Regional Dynamics* (s. 21-43). Berlin Heidelberg: Springer.
- Bilbao-Osorio, B., & Rodríguez-Pose, A. (2004). From R&D to Innovation and Economic Growth in the EU. *Growth and Change*, 35(4), s. 434-455.
- Borch, K., Daimer, S., De Roure, D. C., Deketelaere, K., Dimitropoulos, A., Felt, U., & van der Wende, M. (2015). *The Knowledge Future: Intelligent policy choices for Europe 2050. Report by an expert group on Foresight on Key Long-term Transformations of European systems: Research, Innovation and Higher Education*. Luxembourg: Publications Office of the European Union.
- Boschma, R. A. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39(1), s. 61-74.
- Capello, R. (2008). Space and Theoretical Approaches to Regional Growth. I R. Capello, R. Camagni, B. Chizzolini, & F. U, *Modelling Regional Scenarios for the Enlarged Europe. European Competitiveness and Global Strategies* (s. 13-32). Springer.
- Capello, R. (2009). Spatial Spillovers and Regional Growth: A Cognitive Approach. *European Planning Studies*, 17(5), s. 639-658.
- Capello, R., & Fratesi, U. (2013). Globalization and Endogenous Regional Growth. I R. Crescenzi, & M. percoco, *Geografy, Institutions and Regional Economic Performance* (s. 15-37). Springer-Verlag Berlin Heidelberg.
- Charlot, S., Crescenzi, R., & Musolesi, A. (2014). Econometric modelling of the regional knowledge production funktion in Europe. *Journal of Economic Geography*, s. 1-33.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning an innovation. *Admanistrative Science Quarterly*, 35(1), s. 128-152.
- Cook, S. J., Hays, J. C., & Franzese, R. J. (2015). *Model Specification and Spatial Interdependence*. Hentet 11. 07 2017 fra School of Arts and Science. University of Rochester.:
http://www.sas.rochester.edu/psc/polmeth/papers/Cook_Hays_Franzese.pdf

- Cooke, P., & Leydesdorff, L. (2006). Regional Development in the Knowledge-Based Economy: The Construction of Advantage. *The Journal of Technology Transfer*, 31(1), s. 5-15.
- Coutinho, A. (2013). *PORTUGAL: SCIENCE FRIENDLY DESPITE RELATIVE RESEARCH IMMATURITY*. Hentet fra Euro Scientist: <http://www.euroscientist.com/portugal-science-friendly-despite-relative-research-immaturity/>
- Cowan, R., David, P. A., & Foray, D. (2000). The Explicit Economics of Knowledge Codification and Tacitness. *Industrial and Corporate Change*, 9(2), s. 211-253.
- Cuaresma, J. C., Doppelhofer, G., & Feldkircher. (2012). The Determinants of Economic Growth in European Regions. *Regional Studies*, 48(1), s. 44-67.
- Döring, T., & Schnellenbach, J. (2006). What Do We Know about Geographical Knowledge Spillovers and Regional Growth? A Survey of the Literature. *Regional Studies*, 40(3), s. 375-395.
- EU-oplysningen. (2017). *Regionalpolitik og samhørighed*. Hentet 21. Juli 2017 fra <http://www.eu.dk/da/fakta-om-eu/politikker/regionalpolitik>
- Europa-Kommissionen. (2017b). *Noord-Brabant*. Hentet 29. 07 2017 fra Europa-Kommissionen: <https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/base-profile/north-brabant>
- Europa-Kommissionen. (Juni 2004). *Oberbayern - Economy*. Hentet 20. 07 2017 fra https://circabc.europa.eu/webdav/CircaBC/ESTAT/regportraits/Information/de21_eco.htm
- Europa-Kommissionen. (15. 03 2016). *Stuttgart*. Hentet 20. 07 2017 fra http://cordis.europa.eu/stuttgart/intro_en.html
- Europa-Kommissionen. (2017a). *Franche-Comté*. Hentet 02. 08 2017 fra Europa-Kommissionen: <https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/base-profile/franche-comt%C3%A9>
- Europa-Kommissionen. (2017c). *Emilia-Romagna*. Hentet 28. 07 2017 fra <https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/base-profile/emilia-romagna>
- European Patent Office. (2016). *How to Get a European Patent. A Guide for Applicants*. European Patent Office.
- Eurostat. (2015). *BNP og regnskaber for husholdningerne på regionalt niveau*. Hentet fra Statistics Explained : http://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:GDP_and_household_accounts_at_regional_level/da
- Eurostat. (2015). *Geographical information and maps*, © EuroGeographics for de administrative grænser. Hentet 18. 06 2017 fra Geodata: <http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts#nuts13>
- Eurostat. (30. 03 2017). *Gross Domestic Product (GDP) at Current Market Prices by NUTS2 Regions*. Hentet 24. 04 2017 fra Eurostat: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10r_2gdp&lang=en

- Eurostat. (2017a). *Regional Statistics by NUTS Classification*. Hentet Marts 2017 fra Eurostat: <http://ec.europa.eu/eurostat/data/database>
- Feldman, M. P., & Avnimelech, G. (2011). Knowledge Spillovers and the Geography of Innovation - Revisited: A 20 years' perspective on the Field on Geography og Innovation. I D. B. Audretsch, O. Falck, S. Heblich, & A. Lederer, *Handbook of Research on Innovation and Entrepreneurship* (s. 150-160). Cheltenham, UK & Northampton, USA: Edward Elgar Publishing.
- Foray, D., & Lundvall, B.-Å. (1998). The Knowledge-Based Economy: From the Economics of Knowledge to the Learning Economy. I T. Siesfeld, J. Cefola, & D. Neef, *The Economic Impact of Knowledge* (s. 115-121). Routhledge.
- Greunz, L. (2003). *The Econommics of Knowledge Spillovers*.
- Karlsson, C., & Gråsjö, U. (2012). Knowledge Flows, Knowledge Externalities and Regional Economic Development. *CSIR. Electronic Working Paper Series*.(5).
- Karlsson, C., Warda, P., & Gråsjö, U. (2012). Knowledge Spillovers in Europe: A Meta-Analysis. *Centre of Excellence for Science and Innovation Studies. Electronic Working Paper Series*.(280).
- LeSage, J. P. (2014). What Scientists Need to Know about Spatial Econometrics. *The Review of Regional Studies*, 44(1), s. 13-32.
- LeSage, J. P., & Fischer, M. M. (2008). Spatial Growth Regressions: Model Specification, Estimation and Interpretation. *Spatial Economic Analysis*, 3(3), s. 275-304.
- LeSage, J. P., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press.
- Liu, Z., Griffin, T., & Kirkpatrick, T. L. (2014). Statistical and Economic Techniques for Site-specific Nematode Management. *Journal of Nematology*, 46(1), s. 12-17.
- Lucas, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1), s. 3-42.
- Lundvall, B.-Å., & Johnson, B. (December 1994). The Learning Economy. *Journal of Industry Studies*, 1(2), s. 23-42.
- Monfort, P. (2009). *Regional Convergence, Growth and Interpersonal Inequalities across EU*. Report Working Paper, European Commission.
- Morgan, K. (2004). The Exaggerated Death of Geography: Learning, Proximity and Territorial Innovation Systems. *Journal of Economic Geography*, 4(1), s. 3-21.
- OECD. (1996). *The Knowledge-Based Economy*. OECD.
- OECD. (2009a). *How Regions Grow: Trends and Analysis*. Organization for Economic Co-operation and Development. France: OECD Publishing.
- OECD. (2009b). *Regions matter. Economic Recovery, Innovation and Sustainable Growth*. OECD.
- OECD. (2009c). *Patent Statistics Manual*. Organisation for Economic Co-operation and Development.

- OHIM, & EPO. (2013). *Intellectual Property Rights Intensive Industries: Contribution to Economic Performance and Employment in the European Union*. OHIM & EPO.
- Petkova, R. (2009). German regions lead european R&D. *Statistics in focus*. Eurostat. Hentet fra <http://www.uni-mannheim.de/edz/pdf/statinf/09/KS-SF-09-035-EN.pdf>
- Polanyi, M. (1966). *The Tacit Dimension*. University of Chicago Press.
- Robusto, C. C. (1957). The Cosine-Haversine Formula. *The American Mathematical Monthly*, 64(1), s. 38-40.
- Rodriguez-Pose, A., & Crescenzi, R. (2008). Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. *Regional Studies*, 42(1), s. 51-67.
- Rodríguez-Pose, A., & Crescenzi, R. (2008). Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. *Regional Studies*, 42(1), s. 51-67.
- Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), s. 71-102.
- Ryan, S., & O'Connor, R. V. (n.d). *Development of a Team Measure for Tacit Knowledge in Software Development Teams* .
- Sedgley, N., & Elmslie, B. (April 2004). The Geographic Concentration of Knowledge: Scale, Agglomeration, and Congestion in Innovation Across U.S. States. *international regional science review*, 27(2), s. 111-137.
- Smith, K. (2000). *What is the "Knowledge Economy?" Knowledge-Intensive industries and Distributed Knowledge bases*. Working paper . Hentet 01. 08 2017
- Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1), s. 65-94.
- Storper, M., & Venables, A. J. (2004). Buzz: face to face contact and the urban economy. *Journal of Economic Geography*, 4(4), s. 351-370.
- Verspagen, B. (1991). A New Empirical Approach to catching up or Falling Behind. *Structural Change and Economic Dynamics*, 2(2), s. 359-380.
- Wooldridge, J. M. (2013). *Introductory Econometrics* (1 udg.). Cengage Learning.

Bilag 1

Tabel over regioner der er medtaget i stikprøven.

Landekode	Regionskode	Regionsnavn
AT	AT11	Burgenland (AT)
AT	AT12	Lower Austria
AT	AT13	Vienna
AT	AT21	Carinthia
AT	AT22	Styria
AT	AT31	Upper Austria
AT	AT32	Salzburg
AT	AT33	Tyrol
BE	BE10	Bruxelles-Cap./Brussels Hoofdstedlijk
BE	BE21	Antwerp
BE	BE22	Limburg (B)
BE	BE23	East Flanders
BE	BE24	Flemish Brabant
BE	BE25	West Flanders
BE	BE31	Walloon Brabant
BE	BE32	Hainaut
BE	BE33	Liège
BE	BE34	Luxembourg (B)
BE	BE35	Namur
FR	FR10	Ile-de-France
FR	FR21	Champagne-Ardenne
FR	FR22	Picardy
FR	FR23	Upper Normandy
FR	FR24	Centre-Val de Loire
FR	FR25	Lower Normandy
FR	FR26	Burgundy
FR	FR30	Nord-Pas-de-Calais
FR	FR41	Lorraine
FR	FR42	Alsace
FR	FR43	Franche-Comté
FR	FR51	Pays de la Loire
FR	FR52	Brittany
FR	FR53	Poitou-Charentes
FR	FR61	Aquitaine
FR	FR62	Midi-Pyrénées
FR	FR63	Limousin
FR	FR71	Rhône-Alpes
FR	FR72	Auvergne
FR	FR81	Languedoc-Roussillon
FR	FR82	Provence-Alpes-Côte d'Azur
DE	DE11	Stuttgart
DE	DE12	Karlsruhe

DE	DE13	Freiburg
DE	DE14	Tubingen
DE	DE21	Oberbayern
DE	DE24	Oberfranken
DE	DE25	Mittelfranken
DE	DE26	Unterfranken
DE	DE27	Schwaben
DE	DE30	Berlin
DE	DE40	Brandenburg
DE	DE50	Bremen
DE	DE60	Hamburg
DE	DE71	Darmstadt
DE	DE72	Giessen
DE	DE73	Kassel
DE	DE80	Mecklenburg-Vorpommern
DE	DE91	Braunschweig
DE	DE92	Hannover
DE	DE93	Luneburg
DE	DE94	Weser-Ems
DE	DEA1	Dusseldorf
DE	DEA2	Koln
DE	DEA3	Munster
DE	DEA4	Detmold
DE	DEA5	Arnsberg
DE	DEB1	Koblenz
DE	DEB2	Trier
DE	DEB3	Rheinhessen-Pfalz
DE	DED2	Dresden
DE	DEE0	Sachsen-Anhalt
DE	DEF0	Schleswig-Holstein
DE	DEG0	Thuringen
IR	IE01	Border, Midland and Western
IR	IE02	Southern and Eastern
IT	ITC1	Piedmont
IT	ITC3	Liguria
IT	ITC4	Lombardy
IT	ITF1	Abruzzo
IT	ITF2	Molise
IT	ITF3	Campania
IT	ITF4	Apulia
IT	ITF5	Basilicata
IT	ITF6	Calabria
IT	ITG1	Sicily
IT	ITH1	Province of Bolzano-Bozen
IT	ITH2	Province of Trento
IT	ITH3	Veneto

IT	ITH4	Friuli-Venezia Giulia
IT	ITH5	Emilia-Romagna
IT	ITI1	Tuscany
IT	ITI2	Umbria
IT	ITI3	Marche
IT	ITI4	Lazio
LU	LU00	Luxembourg
NL	NL11	Groningen
NL	NL12	Friesland
NL	NL13	Drenthe
NL	NL21	Overijssel
NL	NL22	Gelderland
NL	NL23	Flevoland
NL	NL31	Utrecht
NL	NL32	Noord-Holland
NL	NL33	Zuid-Holland
NL	NL34	Zeeland
NL	NL41	Noord-Brabant
NL	NL42	Limburg (NL)
PT	PT11	North (PT)
PT	PT16	Central Portugal
PT	PT17	Lisbon
PT	PT18	Alentejo
ES	ES11	Galicia
ES	ES12	Asturias
ES	ES13	Cantabria
ES	ES21	Basque Country
ES	ES22	Navarra
ES	ES23	La Rioja
ES	ES24	Aragon
ES	ES30	Madrid
ES	ES41	Castile and León
ES	ES42	Castile-La Mancha
ES	ES43	Extremadura
ES	ES51	Catalonia
ES	ES52	Valencia
ES	ES53	Balearic Islands
ES	ES61	Andalusia
ES	ES62	Murcia
SE	SE11	Stockholm
SE	SE12	East Middle Sweden
SE	SE21	Småland with Islands
SE	SE22	South Sweden
SE	SE23	West Sweden
SE	SE31	North Middle Sweden
SE	SE32	Central Norrland

UK	UKC2	Tees Valley and Durham
UK	UKD1	Northumberland and Tyne and Wear
UK	UKD3	Cumbria
UK	UKD4	Greater Manchester
UK	UKD5	Lancashire
UK	UKE1	East Yorkshire and Northern Lincolnshire
UK	UKE2	North Yorkshire
UK	UKE3	South Yorkshire
UK	UKE4	West Yorkshire
UK	UKF1	Derbyshire and Nottinghamshire
UK	UKF2	Leicestershire, Rutland and Northamptonshire
UK	UKF3	Lincolnshire
UK	UKG1	Herefordshire, Worcestershire and Warwickshire
UK	UKG2	Shropshire and Staffordshire
UK	UKG3	West Midlands
UK	UKH1	East Anglia
UK	UKH2	Bedfordshire and Hertfordshire
UK	UKH3	Essex
UK	UKI	London
UK	UKJ1	Berkshire, Buckinghamshire and Oxfordshire
UK	UKJ2	Surrey, East and West Sussex
UK	UKJ3	Hampshire and Isle of Wight
UK	UKJ4	Kent
UK	UKK1	Gloucestershire, Wiltshire and Bristol/Bath area
UK	UKK2	Dorset and Somerset
UK	UKK3	Cornwall and Isles of Scilly
UK	UKK4	Devon
UK	UKL1	West Wales and The Valleys
UK	UKL2	East Wales
UK	UKM2	Eastern Scotland
UK	UKM3	South Western Scotland
UK	UKM5	North Eastern Scotland
UK	UKM6	Highlands and Islands

Tabel over udeladte regioner.

Landekode	Regionskode	Regionsnavn
AT	AT34	Vorarlberg
DK	DK01	Capital (DK)
DK	DK02	Zealand
DK	DK03	Southern Denmark
DK	DK04	Central Jutland
DK	DK05	Northern Jutland

FI	FI19	Western Finland
FI	FI1B	Helsinki-Uusimaa
FI	FI1C	Southern Finland
FI	FI1D	Eastern and Northern Finland
FI	FI20	Aaland
FR	FR83	Corsica
DE	DE22	Niederbayern
DE	DE23	Oberpfalz
DE	DEC0	Saarland
DE	DED4	Chemnitz
DE	DED5	Leipzig
EL	EL51	Anatoliki Makedonia - Thraki
EL	EL52	Kentriki Makedonia
EL	EL53	Dytiki Makedonia
EL	EL54	Thessalia
EL	EL61	Ipeiros
EL	EL62	Ionia Nisia
EL	EL63	Dytiki Ellada
EL	EL64	Stereia Ellada
EL	EL65	Peloponnisos
EL	EL30	Attiki
EL	EL41	Voreio Aigaio
EL	EL42	Notio Aigaio
EL	EL43	Kriti
IT	ITC2	Aosta Valley
IT	ITG2	Sardinia
PT	PT15	Algarve
PT	PT20	Azores (PT)
PT	PT30	Madeira (PT)
ES	ES63	Ceuta
ES	ES64	Melilla
ES	ES70	Canary Islands
SE	SE33	Upper Norrland
UK	UKD6	Cheshire
UK	UKD7	Merseyside
UK	UKN0	Northern Ireland (UK)

Bilag 2

Resultater for SDEM for intra og internationale regioner.

Tabel 18 Estimationsresultater for SDEM

	Intra	Inter
	SDEM, W_1	SDEM, W_2
hrst02	0,00052627 (0,00095490)	0,000072806 (0,0014796)
tPAT02	0,000050331 (0,000048062)	0,00021456 ** (0,000066198)
dcommu	-0,014489 (0,011196)	-0,018777 (0,019345)
Lag hrst02	0,0024097 (0,0022887)	-0,0017767 (0,0020249)
Lag tPAT02	0,00024347 (0,00014442)	0,000022851 (0,00014150)
Lag dcommu	0,071326 * (0,033195)	0,048023 (0,043365)
Lambda	0,7466	0,021977
AIC	-452,05	-185
AIC lm	-365,36	-186,97

Kilde: Egne beregninger.

Resultater for robusthedstjek, delstikprøver:

Top-25: Direkte og indirekte effekter model 2 (W_1)

Variable	Estimater	P-værdi
hrst02		
Direkte	0,0001110541	0,94791
Indirekte	0,0024001840	0,63805
Total	0,002511238	0,66727
tPAT02		
Direkte	0,0001102816	0,13694
Indirekte	0,0005236213	0,18566
Total	0,000633903	0,12429
dcommu		
Direkte	0,0670586322	0,12157
Indirekte	-0,0215985513	0,91199
Total	0,045460081	0,94146

Kilde: Egne beregninger