

# 1 Spørgsmål til Asger

- Er der noget galt i at bruge en Cobb Douglas produktionsfunktion til at estimere en eksternalitet?
- Produktionsfunktion - translog vs. CF?
- Der er helt sikkert et positivt bias i vores model. Er det tilfredsstillende blot at skrive sig ud af det, eller skal der korrikeres på en eller anden måde. Bemærk, at vi kun har tilgang til private byerhverv.
- Hvad skal vores policy være?
  - a. Planloven: placering af virksomheder, bystørrelsen, 3D-byer ...
  - b. Klyngepolitik. Hvilke og hvor store er effekterne af disse?
  - c. Government-bashing...
  - d. Hvor meget skal artiklen/resume være en opsummering af opgaven? Og hvor meget skal det være et forsøg på at få noget i avisen? De to kan være vanskelige helt at samstemme.
    - \* Vores umiddelbare ide ligger i at sige, at regeringen måske er lidt i konflikt med sig selv, når den på den ene side prøver at fremelske klyngestrategier (agglomerationer) i et videns og forskningsøjemed, mens den sender videnstunge arbejdspladser ud i provinsen ved udflytning af statslige arbejdspladser.

Hermed en analyse og vurdering af produktivitetsudfordringer særligt i Region sjælland

## 2 Indledning

Vi er interesserede i at undersøge, om der kunne være uudtømte effekt ved fra urbanisering i region Hovedstaden set i forhold til region Sjælland.

## 2.1 Motivation

Der er store forskelle i timeproduktiviteten mellem region Sjælland og region Hovedstaden. En del af disse forskelle skyldes ganske givet branchesammensætning, hvor mere produktive brancher fylder mere i region Hovedstaden end de gør i region Sjælland. Men det kan ikke forklare det hele.

Hvis man i stedet sammenligner samme brancher og dermed ser på, hvad produktivitsforskellen måtte være her, da vil det være muligt at udrede om der gives urbane produktivitetseffekter. Dette kalder vi for *urban learning*.

## 3 Teori

### 3.1 Agglomeration

I dette afsnit vil vi gennemgå den økonomiske teori bag agglomeration og baggrunden for, hvorfor agglomeration påvirker produktivitet. Introduktionen til begrebet er primært baseret på baggrund af bogen "*Economics of Agglomeration*" (2013) af Fujita og Thisse.

### 3.2 Definitioner og introduktion

Agglomerationsøkonomi betegner den eksternalitet, der opstår, når økonomiske personer eller virksomheder geografisk placeres tættere på hinanden.

Agglomeration er indenfor den økonomiske litteratur et relativt nyt aspekt. Især inddrages agglomeration nu i højere grad i estimeringer af eksterne økonomiske effekter af transportomkostninger, hvor den positive eksternalitet ved at samlokalisere agenter indtil for nyligt ikke har været tilstrækkeligt inddraget i cost-benefit analyser. Motivationen for at inddrage agglomeration i sådanne analyser er, at investeringer i infrastruktur sænker den effektive distance mellem økonomiske agenter, hvilket både har interne og eksterne effekter. De interne effekter er fx mindre rej-

setid, færre transportomkostninger, mens de eksterne effekter ved agglomeration er gennem videnskudveksling, større innovation, større specialisering. Relevansen for disse eksterne effekter i en vores sammenhæng er, at de medfører en forøgelse af produktiviteten.

Definitionen på agglomerationen er derfor den paraply, der overlapper alle effekter ved geografisk at samle økonomiske agenter (CITATION). Disse effekter kan både være positive – som eksempelvis større innovation og vidensspillover – mens øget trængsel ved for mange personer på samme sted indgår negativt.

**Et større og tættere arbejdsmarked** som betyder bedre matching mellem arbejdsudbuddet og virksomhedens efterspørgsel af arbejdskraft. Endvidere tillader et stor arbejdsmarked også mere specialisering af arbejdsudbuddet. Dette sker både fordi arbejdskraften har mulighed for at akkumulere human kapital, men også fordi arbejderne har mulighed for at drage nytte af hinandens human kapital.

**Fælles brug af leverandører og infrastruktur.** Virksomheder kan drage nytte af flere om at efterspørge mellemvare eller infrastruktur som kræver store faste omkostning. Dette kunne eksempelvis være en havn, hvor en stor efterspørgsel gør det muligt og rentabelt for en leverandør at foretage investering i nye, og mere effektive, terminaler. På samme måde kan producerenter mindske diversiteten af deres produktporteføljer, og dermed få produktivtetsforbedring gennem specialisering og skala økonomi.

**Produktivtetsforbedring gennem vidensdeling.** Spillover af viden kan ske under formelle og uformelle rammer. De formelle rammer kan dannes gennem fx konferencer, forskningsnetværk, konsulenter eller samarbejde virksomheder imellem. De uformelle rammer sker gennem arbejdskraften. Arbejdskraften kan få ny viden mellem netværk, eller når medarbejdere skifter job, således at en ny virksomhed kan drage nytte af medarbejdernes viden og læring. [6]

Effekterne ved agglomeration bliver ofte opdelt i to teoriske klasser: *lokaliseringsøkonomi* og *urbaniseringsøkonomi*. Lokaliseringsøkonomi er et begreb der bruges når koncentration af ensartet industrier er stor. Ensartet virksomheder som er placeret geografisk tæt på hinanden kan drage fordel af teknologiske fremskridt hos hinanden, således at en klynge af virksomheder specialiseres og dermed får

en komparativfordel i særlige produktioner. Dette gør virksomhederne mere produktive og kan forklare produktivitetsvækst i industrielle distrikter. Det andet begreb, urbaniseringsøkonomi, handler i højere grad om de positive effekter som opstår når virksomheder og arbejdskraften koncentrerer sig geografisk, men begge karakteriseret som alsidige. Alsidigheden betyder at brancher kan drage nytte af hinandens teknologiske fremskridt, general markedsstørrelser og fælles infrastrukturen. Studier som beskæftiger sig med de to effekter separat peger på, at effekterne ved lokaliseringøkonomi typisk er størst, mens at videndrivende service virksomheder oplever positive effekter ved urbanisering [4]. I denne opgave vil vi ikke skelne mellem disse to typer for agglomeration, men som et begreb der indeholde effekter både fra lokalisering og urbanisering.

Hvis agglomerationen fremmes, altså at afstanden i økonomien mindskes, kan der samtidigt opstå problemer. Det første problem kunne fx. opstå hvis en by oplever stor tilflytning, og at der samtidigt ikke investeres proportionelt i infrastrukturen, så risiker man at opleve trængsel i området. Trængsel koster tid, længere rejsetider betyder en øget omkostning for virksomheder og gør derfor produktionen dyrere. Omkostningen kan være direkte, såsom løn til chauffører, sælgere og kørende konsulenter. Trængsel kan også skabe et behov for mere kapital i form af flere transportmidler, fordi fremkommeligheden på tværs af en by bliver svær. Indirekte kan også tænke på, at pendlende medarbejdere får en højere reservationsløn pga. rejsetiden, og der gør matching mellem virksomheder og medarbejdere svær [6]. Et andet problem kan være i forbindelse med den forurening som opstår ved urbanisering. OECD påpeger at væsentlige samfundsøkonomiske omkostninger ved forurening. Så hvis ikke der foretages politisk initiativer til at mindske forureningen vil samfundet få øget omkostning eksempelvis gennem fald i produktiviteten pga. dårlig helbred, højere dødelighed, øget sundhedsinvestering og faldene udbytte ved landbrug [2].

Agglomeration kan også øges ved bedre infrastruktur for på den måde at mindske den fysiske afstand mellem byer og mennesker

lokaliseringøkonomi og urbanisering, forskel mellem de to. Vores mål adskiller ikke de to. I meta-analysen skriver de at det ikke gør den store forskel overordnet

### 3.3 Produktivitetskommissionen

### 3.4 Debatindlæg fra CE

## 4 Metode

### 4.1 Måling af agglomerationseffekter med effektiv tæthed

Der findes flere fremgangsmetoder når agglomerationseffekter skal estimeres. Fælles for metoderne er, at danne et mål som skal repræsentere skalaen af den økonomiske aktivitet i en geografisk kontekst. Når parameteren er bestemt, kan den indgå som variabel i estimationen af en produktionsfunktion, således at vi modellerer og fanger den positive agglomerationsekskternalitet. I meta-analysen [4] gennemgår forfatterne evolutionen af forskellige metoder på dette område.

De tidligste eksempler fokuserede udelukkende på lokale byeffekter, og beror på sig ofte på indbyggerantal. Problemet ved disse mål er, at indbyggerantal i et givent område ikke nødvendigvis er et udtryk for den stedsbestemte økonomiske aktivitet, men indbyggerantallet blot vil være en proxy for bystørrelser - her kan man tænke på byer, hvor arbejdsstyrken pendler andetsteds for at arbejde. Sidenhen blev beskæftigelsen introduceret som mål. Fordelen ved dette mål er, at den lokale beskæftigelse mere isoleret udtrykker den økonomiske aktivitet, og derved er et bedre mål for at bestemme produktivitetsfordele ved samlokation. For at gøre beskæftigelsen robust overfor størrelsesforskelle områder imellem benyttede forskningen sig af tætheds mål - altså mål såsom beskæftigelsen per  $km^2$ .

Et fælles problem for ovenstående mål er, at der implicit i målene er en antagelse om afgrænsede markeder, områder imellem. Dette betyder, at områdernes relative afstand ingen betydning har, og områderne kan ikke drage nytte af hinandens markedede. Denne antagelse umuliggør simpelthen spillover effekter mellem geografiske områder. Spillover effekter forekommer gennem tre kanaler, som beskrevet tidligere; større arbejdsmarkeder mht. matching, vidensdeling og/eller specialise-

Dette er tilfældet i studiet [1], som vores analyse tager udgangspunkt i. Her introducere Graham en parameter, effektiv tæthed, som måler den økonomiske aktivitet i et område, ift. beskæftigelsestætheden, og områdets relative geografisk afstand til andre økonomiske områder ift. deres repsektive beskæftigelsestætheden. Med andre ord indfanger parameteren effekt ved den relative afstanden til andre potentielle markeder, altså forsvinder den implicitte antagelse om afgrænsning områderne i mellem, og målet tillader en afsmittende effekt ved agglomeration. To nærliggende områder med høj beskæftigelsestæthed for altså lov til at drage fordel af hinanden, men samtidigt er den effektive tæthed aftagende i afstand - desto længere to områder er fra hinanden, desto mindre påvirkning.

Figure 4 consists of two maps of Denmark, labeled (a) and (b). Map (a) is titled 'Beskæftigelsestætheden' and shows the employment density in Ringsted Kommune. It includes a map of Denmark with a red circle around Ringsted Kommune, labeled 'Ringsted Kommune'. A red line indicates the 'Radius af Ringsted Kommune' (Radius of Ringsted Kommune). A red arrow points to the 'Beskæftigelsen i Ringsted Kommune' (Employment in Ringsted Kommune). Map (b) is titled 'Potentielle marked' and shows the potential market area for Ringsted Kommune. It includes a map of Denmark with a red circle around Ringsted Kommune, labeled 'Ringsted' with coordinates (55,44295;11,79137). A red arrow points to the 'Slagelse' area, labeled 'Slagelse' with coordinates (55,40369;11,35526). A red arrow also points to the 'Ringsted' area, labeled 'Ringsted' with coordinates (55,44295;11,79137). A red arrow indicates the distance of 27,9 km between the two areas.

6

Til vores analyse har vi valgt, at konstruere parameteren effektiv tæthed,  $ED$ , i tråd med Grahams metode jf. [1]. Fremgangsmetoden kan med fordel forklares med et eksempel som illustreret i figur 1. Vi tager de enkelte kommuner, her Ringsted Kommune, og benytter antal af registerbaseret beskæftigelse efter arbejdssted<sup>1</sup> til skalaen af den økonomiske aktivitet. For at udregne tætheden af beskæftigelsen i Ringsted Kommune, finder vi forholdet mellem beskæftigelsen i kommunen og kommunens størrelse målt på radius kilometer, det betyder at vi implicit antager at kommunerne er cirkelformet. Når vi kender arealet på kommunerne, kan vi blot udregne radius ved hjælp af formlen for radiusen af en cirkel. Denne antagelse viser sig nyttig, når vi skal udregne den relative anstand kommunerne og deres økonomiske aktivitet i mellem. Derefter skal vi bestemme midtpunktet alle kommunerne, for så at udregne beskæftigelsen per afstandskilometer til alle de 97 andre kommuners midtpunkt.

$ED$  måler altså den agglomeration som hver virksomhed oplever givet deres placering i Danmark. Første led er selve agglomerationen i hjemkommunen,  $j$ , mens andet led er den agglomerationseffekt virksomhederne oplever fra resten af landets kommuner som er faldende i afstand. Virksomhederne i Ringsted Kommune påvirkes altså positivt af en høj beskæftigelse i hjemmekommune, samtidigt påvirkes virksomhederne relativt mere af en høj beskæftigelse i nærliggende kommuner, såsom Slagelse, forhold til jyske kommuner. En væsentlig afgrænsning i vores mål er, at vi ikke tilader agglomerationseffekter på tværs af landegrænser. Det betyder at byer som Flensborg og Malmø ikke indgår i analysen, men pendlerne som rejser på tværs af landene for at arbejde indgår i beskæftigelsen. Hvis agglomerationseffekter skal indgå i evalueringen af national politik initiativer er tabet ved denne forsimpelse ikke stor - hvis derimod fordelene ved en Øresundsbro eller Fermenforbindelse skal evalueres kan analysen med fordel udvides. Den effektive tæthed kan opskrives på følgende vis:

$$ED_t^j = \frac{L_t^j}{Radius_j} + \sum_{k=1}^{k \neq j} \frac{L_t^k}{d_{kj}}, \quad Radius_j = \sqrt{\frac{A_j}{\pi}} \quad (1)$$

hvor  $d_{kj}$  er afstanden mellem kommune  $k$  og  $j$ .

---

<sup>1</sup>Registerbaseret arbejdsstyrke, beskæftigelse, tabel RAS301, Danmarks Statistik

For at finde afstandene mellem alle 98 kommuner i Danmark, benytter vi en såkaldt API. En API gør det muligt at automatisere forespørgelser til en server, i dette tilfælde til servicen Google Maps gennem programmet R. Ved at sende navnet på kommune returnerer Google Maps en lokation i form af koordinat i breddegrader og længdegrader. Vi antager at denne lokation er det approksimative økonomiske midtpunkt af en given kommune.

Dernæst udregner vi afstanden mellem to kommuners koordinater vha. Haversine-formel<sup>2</sup>. Haversine-formel beregner den korteste afstanden mellem to punkter på en sfære, hvilket netop giver fulgfulgt mellem to kommunernes midtpunkter. Således kan vi udregne afstandene på kryds og tvær af alle kommuner i en  $98 \times 98$  symmetrisk matrice. Her er det vigtigt at pointere, at vores afstandsmål ikke tager hensyn til den faktiske kørselsafstand mellem to kommuner. Dette betyder at tolkning af parameteren er den procentvise produktivitetsgevinsten hvis afstanden mellem virksomheder/arbejdskraften mindske med en procent.

Vi udregner  $ED$  for perioden 2008-2015<sup>3</sup>. Hvis vi betragte gennemsnittet af variabelen i løbet af perioden, ser vi at de kommuner som er placeret i hovedstadsområdet har den højste værdi. Dette skyldes den høje tæthed af beskæftigelse i København og Frederiksberg. Frederiksberg, Rødovre og København har de tre højste værdi, efterfulgt af de andre kommuner i Storkøbenhavn. Yderligere se vi nogle af de større midtsjællandske kommuner også drager fordel af deres relative afstand til hovedstadsområdet. Aarhus er placeret som nummer 32, efterfulgt af nogle nabokommuner såsom Skanderborg og Odder. Fyns og midtjyllands kommuner findes midt på listen. Den lave ende af listen er kendetegnet ved kommuner, vis placering ofte betegnes som værende i den rådne banan. Det er kommuner som er placeret tæt på den jyske vestkyst. Dette er kommuner såsom Frederikshavn, Hjørring Thisted og Lemvig. I bunden af listen finder vi Bornholm og Læsø som 'staffes' for både at have en tynd beskæftigelsestæthed og deres geografiske placering.

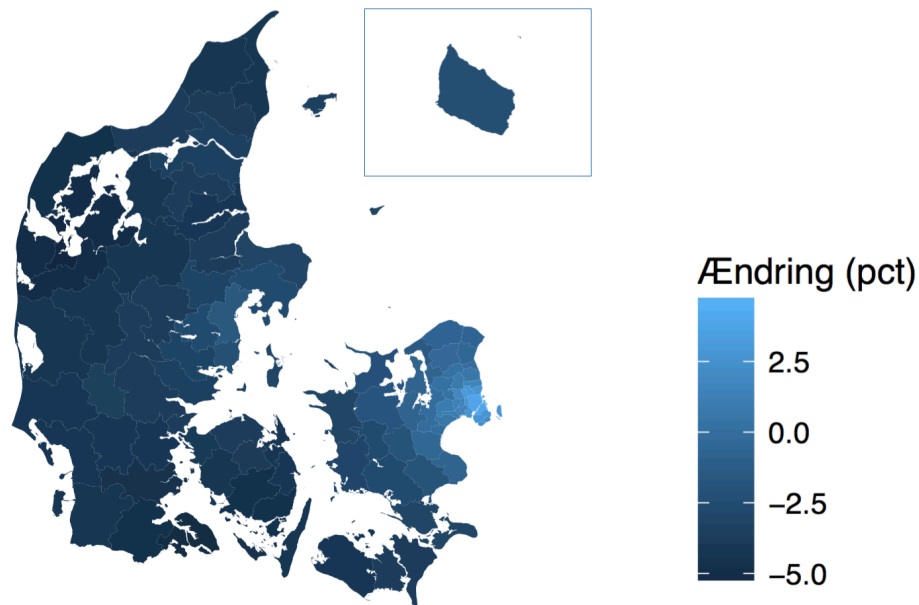
Hvis vi betragter ændringen i effektiv tæthed fra 2008 og 2015 tegner der sig en klart billed af urbanisering omkring hovedstaden. Dette er illustreret på Danmarks-kortet i figur 2. Region Hovedstaden er som en den eneste region, hvor den effektive

---

<sup>2</sup>Skriv en lille historie

<sup>3</sup>En oversigt over effektiv tæthed i perioden 2008-2015 for alle 98 kommuner findes i Appendix





Figur 2: Ændring i effektiv tæthed i Danmark 2008-2015, pct.

tæthed har udviklet sig positivt. Aarhus og de omkringliggende kommuner er næsten uændret i perioden, altså neutral eller svag mindsning af effektiv tæthed. De kommuner hvor udviklingen har været mest negativ på tværs af den 'rådne banan', altså Midt og Vestjylland, Sønderjylland, Fyn og Sydsjælland.

## 4.2 Metode

For at kunne måle agglomerationens effekt for danske virksomheder estimerer vi en *agglomerationselasticitet* ved at estimere en produktionsfunktion for virksomheder i private byerhverv. Derved er det muligt at vurdere indenfor hvilke områder agglomeration betyder mest relativt til andre brancher – og i givet fald hvor meget. På den måde kan vi estimere, hvor meget agglomeration betyder for virksomhedernes værditilvækst, eller med andre ord: Hvor meget af den generede værditilvækst i

danske virksomheder skyldes arbejdskraften efficiente densitet?

## 5 Estimation af produktionsfunktionen

For at kunne svare på dette har vi estimeret en produktionsfunktion for virksomhederne – fordelt efter forskellige branchegrupperinger. Følgende afsnit forklarer metoden, vi har valgt til at bestemme effekten af agglomeration.

Vi antager, at virksomhedernes produktion følger en Cobb-Douglas funktion.

$$Y_{it} = A_t K_{it}^{\beta_k} L_{it}^{\beta_l} ED_{it}^{\beta_{ed}} \quad (2)$$

$ED_{it}$  er arbejdskraftens *efficiente densitet*, der for den enkelte virksomhed indgår som en eksternalitet. Transformeret til logaritmisk form er virksomhedens produktion derfor givet ved følgende relation:

$$y_{it} = \alpha_t + \beta_l l_{it} + \beta_{ed} ed_{it} + \beta_k k_{it} + \epsilon_{it} \quad (3)$$

hvor  $y$  er virksomhedens værditilvækst,  $l_{it}$  er virksomhedens input af arbejdskraft og  $k_{it}$  er virksomhedens input af kapital. Målet er således at estimere  $\beta_{ed}$ , der er vores agglomerationselasticitet.

**Hvis vi estimerede (3) ved hjælp af OLS, skulle vi for at få et konsistent resultat kræve, at virksomhedernes brug af fx arbejdskraft skulle være uafhængigt af hinanden.** Men da vi må formode, at virksomheder med høj produktivitet anvender mere arbejdskraft og kapital end andre, siger vi derfor, at modellen udtrykt ved (3) har et endogenitetsproblem.

Litteraturen fremstiller flere metoder til at korrigere for dette; vi har valgt at følge Levihnsen og Petrins metode [3], der benytter virksomhedernes forbrug af mellemvarer i produktionen til at kontrollere for endogenitetsproblemet. I vores model er metoden udvidet med en parameter for agglomerationen, nemlig  $\beta_{ed} ed_{it}$ .

Levihnsen og Petrins metode foregår i to trin [5, p. 115ff.]. Første trin består i at finde et konsistent estimat for de frie variable i modellen, her  $\beta_l$  og  $\beta_{ED}$ . Modellen

er følgende<sup>4</sup>.

$$y_t = \beta_0 + \beta_l l_t + \beta_{ed} ed_t + \beta_k k_t + \beta_t m_t + \omega_t + \eta_t \quad (4)$$

hvor  $m_t$  er virksomhedernes forbrug af mellemvarer. Modellens residualer deles på i to dele,  $\omega_t + \eta_t$ .  $\eta_t$  angiver den del af residualet, der er uafhængigt af kapital, mens  $\omega_t$  er den del af residualet, der skaber det ovennævnte endogenitetsproblem. Det betyder, at metodens grundlæggende formål er, at korrigere for det bias, der opstår gennem den endogene del af residualet,  $\omega_t$ .

Levinson og Petrin foreslår brug af mellemvarer som proxyvariabel for det ukontrollerede produktivitetsskock i  $\omega_t$ . Intuitionen bag dette resultat er, at for enhver given mængde af kapital i virksomheden, vil virksomheder med høj produktivitet skruer op for produktionen – og dermed også for mellemvareforbruget, fx energi. Denne mellemregning er en måde at korrigere for effekten af  $\omega_t$  på kapital, hvis vi kan påvise en markant positiv sammenhæng mellem  $\omega_t$  og virksomhedernes forbrug af mellemvarer.

Levinsohn og Petrin antager, at  $m_t$  er en funktion af kapital og  $\omega_t$ ,  $m_t = m(k_t, \omega_t)$ , og viser, at denne funktion under visse betingelser er monotont stigende i  $\omega_t$  [3, 322-323]. Når dette er tilfældet, kan funktionen inverteres, og det uobserverede led  $\omega_t$  kan udtrykkes som en funktion af to observerede størrelse, kapital og mellemvarer.

$$\omega_t = \omega(k_t, m_t) \quad (5)$$

$$y_t = \beta_l l_t + \beta_{ed} ED_t + \phi(k_t, m_t) \quad (6)$$

$$\phi(k_t, m_t) \equiv \beta_0 + \beta_k + \omega(k_t, m_t)$$

Som approksimering af  $\omega(k_t, m_t)$  benytter Levinsohn og Petrin sig af et 3. grads polynomium interaktionsled mellem kapital og mellemvarer. Det indsættes i (6),

---

<sup>4</sup>Fodtegnet  $i$  er udeladt af hensyn til overskuelighed, selvom estimationen foretages på panel-data.

hvormed  $\beta_l$  og  $\beta_{ed}$  kan estimeres nonlinearært.

$$y_t = \beta_l l_t + \beta_{ed} E D_t + \delta_0 + \sum_{i=0}^3 \sum_{j=0}^{3-i} d_{ij} k_t^i m_t^j \quad (7)$$

Ved estimering af (7) ved brug af OLS har vi estimatorer for  $\beta_l$  og  $\beta_{ed}$ , der er konsistente. Bemærk at  $\delta_0 \neq \beta_0$  da den sidste gemmer sig i interaktionen i 3. gradspolynomiet i (7). Dette afslutter det første trin.

**Formålet med det andet trin er at finde et konsistent estimat for  $\beta_k$ .** Ved at benytte  $\beta_l$  og  $\beta_{ed}$  fra estimering af (7), kan vi finde et estimat for  $\hat{\phi}$  på følgende måde

$$\hat{\phi} = \hat{v}_t - \hat{\beta}_l l_t - \hat{\beta}_{ed} e d_t \quad (8)$$

$$= \hat{\delta}_0 + \sum_{i=0}^3 \sum_{j=0}^{3-i} \hat{d}_{ij} k_t^i m_t^j - \hat{\beta}_l l_t - \hat{\beta}_{ed} e d_t \quad (9)$$

Fra (9) kan vi dermed finde et estimat for  $\omega_t$ :

$$\hat{\omega} = \hat{\phi} - \beta_k^* k_t \quad (10)$$

hvor  $\beta_k^*$  er et initialt gæt for  $\beta_k$ . Ifølge Levihnson og Petrin kan det antages, at  $\omega_t$  følger en markov proces, som de nonlinenært estimere ved hjælp af et 3. grads polynomium.

$$E[w_t | w_{t-1}] = \gamma_0 + \gamma_1 \omega_{t-1} + \gamma_2 \omega_{t-1}^2 + \gamma_3 \omega_{t-1}^3 + \epsilon_t \quad (11)$$

Et estimat for  $\beta_k$  findes nu gennem en golden section optimerings rutine på følgende kvadrat ved at minimere følgende kvadrat:

$$\min_{\beta_k^*} \left( v_t - \hat{\beta}_l l_t - \hat{\beta}_{ed} e d_t - \beta_k^* k_t - E[w_t | w_{t-1}] \right)^2 \quad (12)$$

$\hat{\beta}_k$  estimeres iterativt igennem (10) og (11), der bidrager til (12). Standardafvigelser og varians-kovariansmatricen benyttet til Wald-test for konstant skalaafkast findes ved hjælp af bootstrapping.

**Bootstrapping** Da (12) skaber et estimat af  $\hat{\beta}_k$  på baggrund af en kompliceret optimerings procedure, er det derfor ikke muligt at udregne analytiske værdier for standardafvigelse. Derfor benytter Levinsohn og Petrin sig af bootstrapping, der er en såkaldt *resampling* metode.

Bootstrapping udregner parameterestimerne for modellen ud fra en stikprøve med tilbagelægning på samme størrelse som datasættet. Når dette gøres tilpas mange gange, kan standardafvigelse for de disse estimer, der laves på baggrund af de gentagne stikprøver, benyttes som konsistente standardafvigelse for vores variable i modellen.

Da vores model er kompliceret og i nogen grad komputeringstung har vi valgt at benyttes 50 iterationer til at konstruere standardafvigelse præsenteret i tabel ???.

## 5.1 Kvalitetskorrigeret arbejdskraft

Når virksomhedernes anvendelse af arbejdskraft indgår i produktionen, er det i et produktivitetsøjemed vigtigt at tage højde for, at medarbejderens produktivitet afhænger af især uddannelse og erfaring.

Derfor har vi et kvalitetskorrigeret arbejdsindeks som arbejdsinput i produktionsfunktionen [7, s. 5]. Denne tager højde for, hvor meget den enkelte virksomhed inddrager af arbejdskraft, så der tages højde for medarbejderens uddannelse og erfaring. Uddannelsesniveau og erfaring kombineres i Danmark Statistiks registre med indberettede regnskabsdata så der for den enkelte virksomhed kan skabes et kvalitetskorrigeret arbejdsinput.

Kvalitet følger DØRS definition og udregnes som [7, s. 5]:

$$\bar{L}_{it} = L_{i0t} + \sum_{f=1}^{F-1} \frac{\bar{w}_{fgt}}{\bar{w}_{0gt}} L_{ift} \quad (13)$$

$f$  er her en gruppering af beskæftigelsen efter kvalitet. Fx er  $f = 0$  en ufaglært medarbejder med under 5 års erfaring,  $f = 1$  er en ufaglært medarbejder med 5-10

års erfaring osv. CHECK!!!!!!  $L_{ift}$  er så antallet af årsværk i virksomheder med kvalitet  $f$ . Dermed er det muligt at korrigere for kvaliteten af arbejdskraften, så en times præsteret nu tæller forskelligt alt efter produktiviteten.

I produktionsfunktionen fra (6) indgår den kvalitetskorrigerede arbejdskraft som  $l_{it} = \log(L_{it})$ .

## 6 Beskrivelse af data

Vi har benyttet Danmarks Statistiks virksomhedsregister som baggrund for estimeringen. Data tilgængeligt for os er i årene 2008 – 2013, hvori vi har 45.751 observationer fordelt på ca. 17.000 cvr numre. Datasættet er ikke et fuldt panel. For en beskrivelse af agglomerationsvariablen, jf. kapital ???.

Virksomhederne er kun indenfor branche i private byerhverv, hvilket vil sige branchegrupperne Industri, Bygge og anlæg, Transport, Hoteller og restauranter, Handel, Information og Kommunikation, Vidensservice og Operationel service.

Der er lavet visse restriktioner på registrene. Blandt andet er virksomheder med færre end fem ansatte.

**Estimation af kapitalapparat** I estimationen af kapitalapparatet har vi anvendt dit og dat som du bør skrive mere om.

- Spørg Lill om metoden

**Antagelser for konsistent estimering af ed**

## 7 Estimation

Vi ønsker at estimere følgende output dataset

$$\ln Y_{it}^{pj} = \alpha_0^p + \ln K_{it} + \ln L_{it} + ED_t^j + \omega_t^p \quad (14)$$

	Alle brancher	Bygge og anlæg	Handel	Info. og komm.	Hotel og rest.	Transport	Videnservice	Op. service
$\beta_t$	0.810*** (0.007)	0.724*** (0.014)	0.893*** (0.011)	0.788*** (0.039)	0.714*** (0.035)	0.721*** (0.038)	0.825*** (0.024)	0.721*** (0.031)
$\beta_{ed}$	0.169*** (0.007)	0.107*** (0.013)	0.137*** (0.011)	0.185*** (0.028)	0.147*** (0.025)	0.124*** (0.029)	0.221*** (0.017)	0.0790** (0.038)
$\beta_k$	0.0663*** (0.007)	0.0911*** (0.022)	0.0548*** (0.017)	0.0826*** (0.027)	0.0604** (0.029)	0.144*** (0.041)	0.0513** (0.023)	0.0723*** (0.020)
N	42,751	5,290	14,022	2,625	1,678	2,776	3,914	2,487
Wald_1	18,080	8,896	22,170	1,297	1,713	0,043	10,130	5,633
p-værdi	0,000	0,003	0,000	0,255	0,191	0,835	0,001	0,018
Wald_2	119,600	47,620	7,928	8,837	25,410	8,614	25,840	32,280
p-værdi	0,000	0,000	0,005	0,003	0,000	0,003	0,000	0,000

Standard errors in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1  
Kilde: Danmarks Statistik og egne beregninger.

Tabel 1: Outputtabel

	Hotel og rest.	Transport	Videnservice	Op. service
$\beta_t$	0.714*** (0.0350)	0.721*** (0.0377)	0.825*** (0.0240)	0.721*** (0.0307)
$\beta_{ed}$	0.147*** (0.0250)	0.124*** (0.0294)	0.221*** (0.0168)	0.0790** (0.0379)
$\beta_k$	0.0604** (0.0289)	0.144*** (0.0405)	0.0513** (0.0227)	0.0723*** (0.0196)
N	1,678	2,776	3,914	2,487
Wald_1	1,713	0,043	10,130	5,633
p-værdi	0,191	0,835	0,001	0,018
Wald_2	25,410	8,614	25,840	32,280
p-værdi	0,000	0,003	0,000	0,000

Standardafvigelser i parantes, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1  
Kilde: Danmarks Statistik og egne beregninger.

Tabel 2: Outputtabel

hvor  $i$  er virksomhedsindeks,  $p$  kommuneindeks,  $t$  angiver tidspunkt i år og  $j$  angiver kommune.

## 8 Praktisk case/vinkel

## 9 Diskussion

Forslag: Kvalitetsjustering af beskæftigelsen i ED og lag værdier af ED

	Industri, samlet	Fødevarer	Kemisk	Maskin	Øvrig industri
$\beta_l$	0.695*** (0.0126)	0.621*** (0.0336)	0.704*** (0.0732)	0.703*** (0.0239)	0.727*** (0.0203)
$\beta_{ed}$	0.0514*** (0.0143)	0.111*** (0.0378)	0.0835 (0.0768)	0.0642** (0.0299)	0.0322 (0.0206)
$\beta_k$	0.117*** (0.0202)	0.116** (0.0572)	0.202 (0.247)	0.144*** (0.0379)	0.101*** (0.0349)
N	9,959	1,274	280	2,069	6,336
Wald <sub>1</sub>	23,760	5,895	0,002	3,015	13,620
p-værdi	0,000	0,015	0,966	0,082	0,000
Wald <sub>2</sub>	59,350	18,770	0,157	17,450	23,230
p-værdi	0,000	0,000	0,692	0,000	0,000

Standardafvigelse i parentes, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1  
Kilde: Danmarks Statistik og egne beregninger.

Tabel 3: Outputtabel

## Konklusion

### Kvalitetsjusteret arbejdskraft

**Levinson og Petrin** Hvad nu hvis jeg skriver nog

Syntaksen for at citere er [4, pp. 211ff.].

## Litteratur

- [1] Daniel J Graham. Agglomeration, productivity and transport investment. *Journal of transport economics and policy (JTEP)*, 41(3):317–343, 2007.
- [2] Zbigniew Klimont and Fabian Wagner IIASA. The economic consequences of air pollution.
- [3] James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341, 2003.



- [4] Patricia C Melo, Daniel J Graham, and Robert B Noland. A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics*, 39(3):332–342, 2009.
- [5] Amil Petrin, Brian P Poi, James Levinsohn, et al. Production function estimation in stata using inputs to control for unobservables. *Stata journal*, 4:113–123, 2004.
- [6] Peter Birch Sørensen, Carl-Johan Dalgaard, Agnete Gjersing, Hans Nikolaisen, Agnete Raaschou-Nielsen, Philipp Schröder, Jan Rose Skaksen, Birgitte Sloth, and Anders Sørensen. *Infrastruktur*. 2014.
- [7] De Økonomiske Råd. Baggrundsnotat: Estimation af tfp og sammenhæng til uddannelse. 2010.