

Dokumentation for produktivitetsanalyse

Produktiviteten er illustreret ved spredningen i produktiviteten på brancher og bruges til udpegningen af brancher med tegn på manglende konkurrence.

Datagrundlag

Kilder

Kilde til afkastningsgraden er Købmandstandens Oplysningsbureau. Købmandstandens Oplysningsbureaus database dækker virksomheder i Danmark, som har indberettet regnskaber til Erhvervs og Selskabsstyrelsen. For de ca. 527.000 virksomheder i databasen, er der oplysninger for op til 6 år. Men ikke alle virksomheder har indrapporteret alle regnskabsoplysningerne, eksempelvis har kun ca. 29.000 virksomheder oplyst omsætningen i et eller flere år. Mange af virksomhederne har ikke indrapporteret andet end adressen. En observation består af regnskabsoplysninger om en virksomhed i et år. I udtræksbasen for 2001 er der i alt 93.974 observationer, det vil sige antallet af virksomheder, hvor afslutningsåret for regnskabet er 2001.

Valideringen af databasen viser, at der er et betragteligt antal fejlobservationer, der er forsøgt korrigeret for fejlobservationerne, men det kan ikke afvises, at der er systematik i fejlene, sådan at bestemte typer af virksomheder enten bliver under- eller overrepræsenterede i databaserne. På samme måde, kan der være systematik i hvilke typer virksomheder, som har valgt ikke at indrapportere alle oplysningerne. Derfor er der risiko for, at opgørelser, der fx bygger på omsætningen, kan være behæftet med systematiske fejl. Dette ændrer resultaterne i forhold til hvis man havde haft adgang til oplysninger fra samtlige virksomheder.

I alt udgår 47.839 ud af 93.974 observationer i Købmandstanden på grund af et eller flere af de nedenstående valideringskriterier. Begge databaser er validerede, så det er sikret for alle observationer at:

Da virksomhederne ofte ikke har givet oplysninger om alle variablene, er der mange observationer, som indeholder en "missing" variabel. Disse observationer kan dog alligevel godt indeholde værdifulde oplysninger i de øvrige variable.

Samme problemer med data gælder for Amadeus-databasen, idet Amadeus er en samling af de tilsvarende databaser i de europæiske lande. For Danmark dækker Amadeus databasen dog kun ca. en tredjedel af Købmandstanden.

Aggregeringsniveau

Spredningen i produktiviteten opgøres på 4-cifret NACE-kode niveau, hvilket bruges til udpegningen af brancher med tegn på manglende konkurrence.

Datausikkerhed

Det vurderes at der er en betydelig usikkerhed forbundet med data.

Metode og beregninger

Metoden til at estimere og beregne spredningen i produktiviteten er inspireret af Erik Strøjer Madsen, Tor Eriksson, Valdemar Smith og Mogens Dilling-Hansen "Produktivitetsforskelle mellem danske virksomheder – afhængighed af brancheforhold" Working Paper 00-4, Department of Economics, Aarhus School of Business, p. 16, 2000.

Produktivitet defineres som total faktor produktiviteten (TFP) på basis af en Cobb-Douglas produktionsfunktion:

$$Y_{it} = TFP_{it} \cdot N_{it}^{\beta_1} \cdot K_{it}^{\beta_2}$$

Tages logaritmer fremkommer den estimerbare form:

$$y_{it} = \alpha_t + \beta_1 n_{it} + \beta_2 k_{it} + \varepsilon_{it}$$

hvor y_{it} er logaritmen til bruttoavancen i år t for virksomhed i , n_{it} er logaritmen til antal ansatte i virksomheden, og k_{it} er logaritmen til kapitalapparatet, hvor virksomhedens samlede aktiver er brugt som mål. Virksomheder, der havde en negativ bruttoavance, er ikke medtaget, da man ikke kan tage logaritmen til et negativt tal. Dette gør, at spredningsmålene er biased i nedadgående retning.

Koefficienterne b_1 og b_2 er outputelasticiteterne med hensyn til hhv. arbejdskraft og kapital og antages konstante over tiden. a_t er logaritmen til TFP og antages at ændre sig over tiden. e_{it} er et restled, som angiver, hvor meget af outputtet, som ikke kan forklares af modellen. Spredningen på e_i viser, hvor stor forskellen på produktiviteten mellem virksomhederne i en branche er.

Modellen er estimeret på data fra Købmandstandens Oplysningsbureau. Modellen tager ikke højde for stigende arbejdsomkostninger, da de forklarende variable, med undtagelse af antal ansatte, er i en form for løbende priser. Der er estimeret særskilte modeller for hver branche.

Cobb-Douglas produktionsfunktionen er en meget simpel model for produktionsfunktionen, og den tager ikke hensyn til inputs, som ikke er arbejdskraft eller kapital, dvs. råvarer mm. Derudover skal fejlledet, som bruges som estimat på logaritmen til den totale faktorproduktivitet, være uafhængig, identisk fordelt med middelværdi 0 og konstant varians. Endvidere benyttes nominelle størrelser, som selvom de bliver trend-renset ved hjælp af a_t , ikke tager hensyn til *forskelle* i prisudvikling mellem brancherne.