# Efficiency benchmarking project B:
# Analogous efficiency measurement model based on Stochastic Frontier Analysis

## Final Report

11.12.2006

**Mikko Syrjänen, Gaia Consulting Oy**
**Peter Bogetoft, Sumicsid AB**
**Per Agrell, Sumicsid AB**

# CONTENTS

# Suomenkielinen tiivistelmä

## Johdanto

Energiamarkkinavirasto (EMV) on käynnistänyt sähkönjakeluverkkoliiketoiminnan hinnoittelun valvontamallin kehityshankkeen vuosien 2008-2011 valvontajaksoa varten. Hankkeen tavoitteena on muun muassa sähkönjakelun tehokkuuden arviointimallin jatkokehittäminen, jakeluverkkoyhtiöiden pitkän tähtäimen tehostamispotentiaalin määrittäminen, yhtiökohtaisten tehostamistavoitteiden asettaminen ja vaihtoehtoisten valvontamallien kannustinvaikutusten analysointi. Valvontamallin kehittämisen näkökulmasta keskeisenä tavoitteena on asettaa jakeluverkkoyhtiöille yhtiökohtainen tehostamistavoite, joka heijastelee yhtiöiden toiminnan tehostamispotentiaalia.

Kehityshankkeen osana EMV on tilannut kaksi kehitysprojektia, jotka tähtäävät tehokkuuden arviointimallien kehittämiseen. Tavoitteena on ollut kehittää tehokkuuden arviointiin menetelmiä, joita voidaan käyttää pohjana asetettaessa yhtiökohtaisia tehostamistavoitteita. Osaprojekti A[1] on tähdännyt DEA-menetelmään pohjautuvan tehokkuuden arviointimallin jatkokehittämiseen ja tehokkuuden arviointimallin hyödyntämiseen liittyvien kysymysten analysointiin. Osaprojekti B on tähdännyt rinnakkaisen tehokkuuden arviointimallin kehittämiseen. Yhtiökohtaisen tehostamistavoitteen asettamisen rinnalla tavoitteena on samalla ollut laajentaa tehokkuuden arviointia huomioimaan laatu ja pääomakustannukset. Mallien kehittämisen liittyvien osaprojektien rinnalla EMV on itse vastannut mallien soveltamiseen liittyvien kysymysten kuten tehokkuuslukujen käyttöön liittyvien kysymysten ratkaisemisesta. Tässä tiivistelmässä esitellään osaprojektin B keskeisiä tuloksia.

Osaprojektin B tavoitteena on ollut kehittää rinnakkainen tehokkuuden arviointimalli, jota voidaan käyttää DEA-mallin tulosten varmennukseen. Kehitystyön lähtökohtana on ollut luoda malli, jolla voidaan ratkaista varsinkin DEA-menetelmään liittyviä estimointiongelmia, kuten ominaisuuksiltaan tai kooltaan poikkeuksellisten yhtiöiden tehokkuuden arviointi sekä DEA-menetelmän herkkyys aineistossa esiintyvälle satunnaisvaihtelulle. Siten tavoitteena on ollut luoda malli, joka on vertailukelpoinen osaprojektissa A kehitettävän DEA-mallin kanssa. Hankkeen lähtökohtana on kuitenkin ollut valita mallin spesifikaatiot itsenäisesti niin, että tulokset ovat mahdollisimman riippumattomia. Tavoitteena ei siten ole ollut tuottaa DEA-menetelmä kanssa identtisiä tuloksia vaan tarjota riippumaton vertailukohta. Kirjallisuuskatsauksen perusteella hankkeen lähestymistavaksi valittiin Stochastic Frontier Analysis (SFA) -menetelmä, joka tarjoaa DEAn ominaisuuksia täydentävän lähestymistavan.

Hankkeen painopisteenä on ollut SFA-menetelmän testaaminen ja mallin konseptuaalinen kehittäminen. Tarkastelussa on siten keskitytty panosten ja tuotosten valitaan nykyisin käytössä olevien tekijöiden joukosta, funktiomuodon valintaan sekä tehokkuuden ja satunnaisvaihtelun jakaumia koskeviin oletuksiin. Tässä vaiheessa mallin yksityiskohtaiset parametrit tai yhtiökohtaiset tehokkuusluvut eivät ole olleet tarkastelun pääasiallisena kohteena. Koska ensisijaisena tavoitteena on ollut tarkastella estimointiin liittyviä ongelmia, panos- ja tuotostekijöiden valinnassa on pitkälti tukeuduttu DEA-mallia varten kerättyyn aineistoon.

---

[1] Osaprojektit A ja B toteutettiin samanaikaisesti. Osaprojektin A tulokset esittelee Honkapuro, Tahvanainen, Viljainen, Lassila, Partanen, Kivikko, Mäkinen, Järventausta (2006).

Mallin kehittämiseksi hankkeessa on arvioitu useita eri panos-tuotos-yhdistelmiä ja funktiomuotoja. Vaihtoehtoisia malleja on arvioitu useista näkökulmista huomioiden niin käsitteellinen selkeys ja johdonmukaisuus, tilastolliset ominaisuudet kuin valvontakäytön ja käytännön näkökulma. Käytetty aineisto on ollut identtinen rinnakkaiseen DEA-menetelmän kehittämiseen tähdänneessä osaprojektissa, mikä mahdollistaa tulosten suoran vertaamisen. Pääasiallisen aineistona on käytetty vuoden 2004 dataa. Mallin ja tulosten stabiiliutta on analysoitu suhteessa vuoden 2003 aineistoon.

Hankkeen lähtökohtana on ollut kehittää DEA-mallille rinnakkainen malli, jota voidaan käyttää tulosten varmentamiseen. Tavoitteena on ollut kehittää malli, jota EMV voi käyttää itsenäisesti tehokkuuslukujen varmentamiseen ja yksittäisten jakeluverkkoyhtiöiden tehostamispotentiaalien arviointiin. Mallin varsinaista käyttötapaa ei kuitenkaan tässä vaiheessa ole vielä päätetty. Lähtökohtana on kuitenkin ollut kehittää SFA-malli, jota voidaan haluttaessa käyttää rinnakkain DEA-mallin kanssa varsinaisessa valvonnassa.

## Stochastic Frontier Analysis -menetelmä

SFA- ja DEA-menetelmä perustuvat yhteiseen lähtökohtaan tehokkaan rintaman eli tehokkaan tuotantofunktion estimoinnista sekä tehokkuuden arvioimisesta suhteessa tähän rintamaan. DEA-menetelmään verrattuna SFA-menetelmän keskeisin ero on siinä, että menetelmä perustuu regressioanalyysin tavoin etukäteen valittavaan funktiomuotoon, joka sovitetaan aineistoon määrittämällä rajattu määrä parametreja. Tämän lisäksi menetelmä huomioi aineistossa olevan satunnaisvaihtelun. Nämä ominaisuudet tuovat mukanaan seuraavia etuja.

- Mahdollisuus testata tilastollista merkitsevyyttä muun muassa panos- ja tuotostekijöiden valinnassa
- Tulosten pienempi herkkyys aineistossa esiintyville virheille ja vaihtelulle
- Tehokkuuden ja satunnaisvaihtelun erottaminen
- Yksittäisten poikkeavien yhtiöidenkään tehokkuusluku ei koskaan perustu vertailuun vain yhtiön itsensä kanssa.

Toisaalta SFA-menetelmän heikkoutena on se, että funktiomuoto sekä tehokkuuden ja satunnaisvaihtelun jakauma joudutaan olettamaan etukäteen.

Tehokkuuden arviointimenetelmän valinnassa joudutaan aina tekemään kompromissi sen suhteen, miten joustavasti kustannusrakenne voidaan arvioida eli miten joustavasti malli seuraa aineistoa ja toisaalta miten herkkä malli on aineistossa olevalle satunnaisvaihtelulle. SFA korostaa mallin kykyä suodattaa aineistossa olevia satunnaisvirheitä. Vastaavasti DEA-menetelmän vahvuutena on joustavuus, jonka ansiosta malli kuvaa hyvin panosten ja tuotosten riippuvuutta aineiston perusteella. Mikäli mallin rakenne on oikein valittu, menetelmät tuottavat eroistaan huolimatta hyvin samankaltaisia tuloksia.

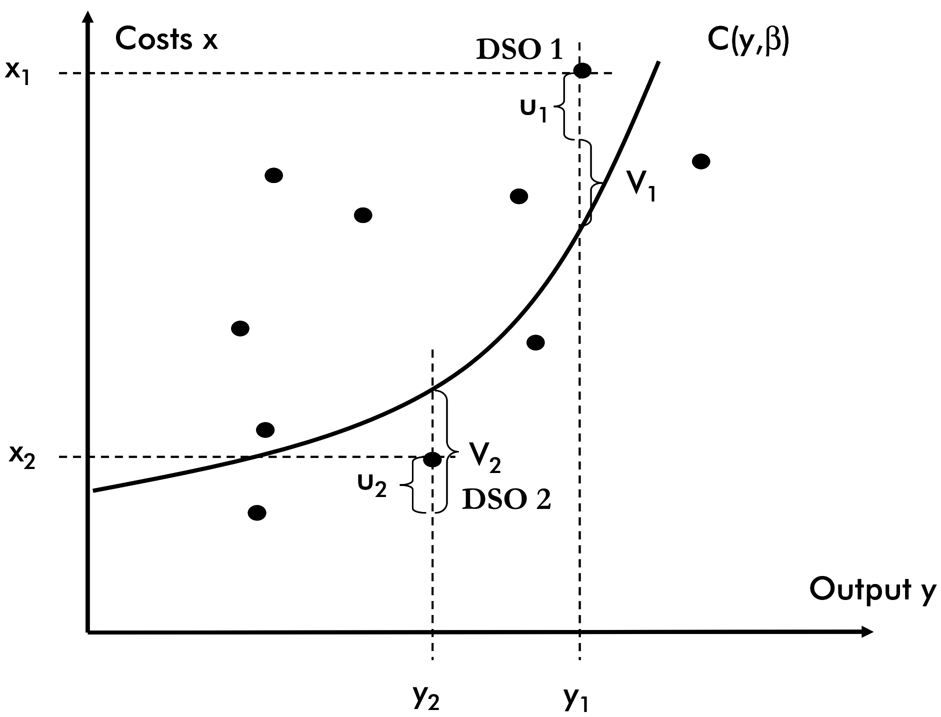
Seuraavassa on tiivistetysti kuvattu SFA-menetelmän ominaisuuksia. Varsinainen englanninkielinen raportti sisältää kattavamman esityksen ja esimerkiksi Coelli et al. (1998) esittävät perusteellisen johdatuksen aiheeseen.

SFA-menetelmän soveltamisessa lähtökohtana on tässä tapauksessa mallintaa jakeluyhtiöiden kustannuksia useiden tuotostekijöiden funktiona. Menetelmän keskeisenä oletuksena on se, että yhtiöt poikkeavat estimoidusta kustannusfunktiosta eli rintamasta niin tehottomuuden kuin satunnaisen vaihtelun takia. Matemaattisesti menetelmän perusajatus voidaan siis esittää seuraavasti.

$x_i = C(y_i) + u_i + v_i$, $i = 1,\ldots,N$

Tässä $x_i$ on yhtiön i havaittu kustannustaso ja $C(y_i)$ on yhtiön i tuotoksia $y_i$ vastaava tehokas kustannustaso. Edelleen $u_i$ on yhtiön i yksilöllinen tehottomuus, ja $v_i$ on yhtiön havaittuun kustannustasoon liittyvä satunnaisvaihtelu. Tässä hankkeessa satunnaisvaihtelun $v_i$ on oletettu olevan normaalijakautunut ja tehottomuuden $u_i$ vastaavasti noudattavan katkaistua normaalijakaumaa, jossa jakauman katkaisukohta estimoidaan aineistosta.

Kuva 1 esittää SFA-menetelmän perusajatuksen graafisesti. Kuvassa yhtenäinen käyrä ($C(y,\beta)$) esittää tehokasta kustannustasoa. Havaitut pisteet eroavat tästä kahden tekijän seurauksena. Satunnaisvaihtelu on merkitty v:llä ja tehottomuus u:lla. Kuvassa on havainnollistettu kahden yhtiön tilannetta. Yhtiön DSO 1 tapauksessa tehokkaan tason ylittävä kustannus aiheutuu positiivisesta satunnaistermistä $v_1$ (esim. satunnainen poikkeava kustannuserä) sekä tehottomuudesta $u_1$. Toisaalta yhtiön DSO 2 tapauksessa yhtiön poikkeuksellisen alhainen kustannustaso johtuu kustannustasossa olevasta satunnaisesta virheestä $v_2$, joka ylittää yhtiön tehottomuuden $u_2$.




**Kuva 1** SFA mallin perusajatus

Teknisesti tehokkuuden ja satunnaisen vaihtelun erottaminen tapahtuu mallin estimoinnin yhteydessä perustuen siihen oletukseen, että satunnaisvaihtelu on symmetrisesti jakautunut ja tehottomuus epäsymmetrisesti. Käytännössä kustannusfunktion parametrit ja jakaumien parametrit on estimoitava samanaikaisesti maximum likelihood -estimoinnilla. Tavoitteena on etsiä ne kustannusfunktion ja jakaumien parametrit, jotka olisivat todennäköisimmin tuottaneet havaitun aineiston.

Edellä kuvatussa kaavassa tehottomuus on esitetty terminä, joka lisätään tehokkaaseen kustannustasoon. Mallia estimoitaessa tehottomuus esitetään siis absoluuttisena euromääräisenä tehottomuutena. Jotta tuloksia voidaan verrata DEA-menetelmään, tämä tehottomuus on tulosten arvioinnissa muutettu suhteelliseksi tehokkuusluvuksi. Yhtiön i tehokkuusluku $E_i$ lasketaan siten seuraavasti: $E_i = C(y_i) / [C(y_i) + u_i]$.

Käytännössä SFA-malli estimointi vaatii erityisen ohjelmiston. Mallin estimointi voidaan tehdä erityisesti tähän tarkoitukseen rakennetuilla ohjelmistoilla kuten LIMDEP[2] ja Frontier[3]. Nämä tarjoavat helpoimman tavan laskea SFA-tuloksia. Toisaalta SFA-malleja voidaan estimoida yleisillä tilasto-ohjelmilla kuten R, S, SAS tai GAMS.

## Rinnakkaisen mallin rakentaminen

Rinnakkaisen SFA-menetelmään perustuvan mallin rakentamisessa testattiin lukuisia panos-tuotos-yhdistelmiä sekä funktiomuotoja. Käytännössä panosten ja tuotosten ja funktiomuodon valintaan on tarkasteltu samanaikaisesti perustuen yhteensä 24 tuotos-panos-yhdistelmään ja neljään perusfunktiomuotoon. Seuraavassa tuloksia tarkastellaan kuitenkin selkeyden vuoksi erikseen panosten ja tuotosten valinnan sekä funktiomuodon valinnan osalta.

### Panos- ja tuotostekijät

Tämän osaprojektin ensisijaisena tavoitteena ei ole ollut tarkastella kattavasti mahdollisia tuotos- ja panostekijöitä. Näihin tekijöihin ei tällä hetkellä ole katsottu liittyvän keskeisiä kehityskysymyksiä. Siten tekijöiden valinnan pohjana oleva joukko on perustunut DEA-menetelmään pohjautuneisiin tutkimuksiin, kuten Korhonen et al. (2000) ja samanaikaisesti käynnissä ollut osaprojekti A. Hankkeessa on kuitenkin testattu useita panos-tuotos-yhdistelmiä.

Panosten osalta hankkeessa on tarkasteltu neljää vaihtoehtoista panostekijää, jotka ovat koostuneet operatiivisista kustannuksista (opex), poistoista (depreciation) ja KAH-arviojen pohjalta määritellyistä keskeytyskustannuksista (interruption costs). Kokeiltujen mallien perusteella kaikki kolme kustannuskomponenttia sisältävä panostekijä toimii mallin estimoinnissa hyvin.

Tuotospuolella lähtökohtana ovat olleet tällä hetkellä käytössä olevan DEA-mallin tuotostekijät. Tuotostekijöiden joukon laajentamista rajoittaa saatavissa olevan aineiston määrittely. Tuotostekijöinä on käytetty siirretyn energian arvoa (value of energy), verkostopituutta (total network length) sekä asiakasmäärää (No. of customers). Näiden lisäksi on tarkasteltu vaihtoehtona verkoston jälleenhankinta-arvon käyttöä verkostopituuden sijaan. Jatkossa tuotosjoukkoa voitaisiin laajentaa esim. jakamalla verkostopituus kaapeli- ja ilmajohtoverkkoon.

Aineiston analyysissä nykyisin DEA-mallissa käytössä oleva tuotosjoukko todettiin tarkasteltujen vaihtoehtojen joukossa hyvin toimivaksi.

### Mallin funktiomuoto

Hankkeessa testattiin useita funktiomuotoja. Seuraavassa on esitetty eri funktiomuotoja koskevat johtopäätökset.

---

[2] Lisätietoa osoitteesta http://www.limdep.com/ (englanniksi)
[3] Frontier-ohjelmisto on saatavissa käyttöön korvauksetta. Lisätietoa osoitteesta http://www.uq.edu.au/economics/cepa/frontier.htm (englanniksi)

Yksinkertaisimpana mallirakenteena testattiin tavallista lineaarista mallia $x_i = b_0 + b_1y_{1i} + b_2y_{2i}$. Sen etuna on yksinkertainen tulkinta, mutta funktiomuoto voi olla liian yksikertainen kuvaamaan todellista kustannusrakennetta. Lineaarisen mallin suurin ongelma liittyy aineistossa esiintyvään heteroskedastisuuteen eli siihen, että euromääräinen poikkeama tehokkaasta kustannustasosta on luonnostaan suurempi suurille yhtiöille. Koska mallin estimoinnissa huomioidaan juuri euromääräinen tehottomuus, lineaarisen mallin käyttöä johtaa erikokoisten yhtiöiden kannalta selvästi epätasapainoiseen tulokseen, eikä sen käyttöä voi suositella.

Toisena mallirakenteena testattiin loglineaarista mallia, jossa aineistosta on otettu logaritmit. Mallin rakenne on siten $\ln x_i = b_0 + b_1\ln y_{1i} + b_2\ln y_{2i}$. Tämä mallirakenne ratkaise heteroskedastusuuteen liittyvän ongelman ja malli toimii tilastollisesti hyvin. Malliin liittyy kuitenkin keskeinen käsitteellinen ongelma, sillä tuotantomahdollisuuksien joukko ei ole konveksi, eli kahden tehokkaalla rintamalla sijaitsevan tuotos-panos-pisteen yhdistelmä ei kuulu tuotantomahdollisuuksien joukkoon. Tästä syystä mallin käyttöä ei voi suositella.

Kolmantena ja joustavimpana funktiomuotona tarkastelussa oli translog malli: $\ln x_i = b_0 + b_1\ln y_{1i} + b_2\ln y_{2i} + 0.5b_{11}(\ln y_{1i})^2 + 0.5b_{22}(\ln y_{2i})^2 + b_{12}\ln y_{1i}\ln y_{2i}$. Tämä funktiomuoto on äärimmäisen joustava ja tarjoaa siten teoreettisesti hyvä lähtökohdan. Mallin tilastollisiin ominaisuuksiin liittyy kuitenkin niin merkittäviä ongelmia, ettei malli ole tässä tapauksessa käyttökelpoinen.

Neljäntenä tarkasteltuna funktiomuotona oli normeerattu lineaarinen malli, joka huomioi aineistossa esiintyvän heteroskedastisuuden. Estimoitu malli oli tässä tapauksessa siis muotoa $x_i/y_{1i}= b_0/y_{1i} + b_1y_{1i}/y_{1i} + b_2y_{2i}/y_{1i}$. Normeerauksen, eli data jakamisen tuotostekijällä, tavoitteena on huomioida se, että euroissa mitattu tehottomuus ja satunnaisvaihtelu riippuvat yhtiön koosta. Käsitteellisesti malli on yhtenevä lineaarisen mallin kanssa, mutta se toimii käytännössä huomattavasti paremmin niin tilastollisesti kuin käytännön näkökulmasta. Tulokset ovat hyvin linjassa DEA-menetelmän tulosten kanssa. Testatuista funktiomuodoista tämä näyttää, mahdollisesta liiallisesti yksinkertaisuudestaan huolimatta, toimivan parhaiten.

## Suositeltava mallirakenne

Eri mallivaihtoehtojen analyysin perusteella hankkeessa päädyttiin suosittelemaan panostekijäksi operatiiviset kulut, poistot ja keskeytyskustannukset sisältävää kokonaiskustannusta. Tuotospuolella suositellaan käytettäväksi nykyisen DEA-mallin tuotoksia eli siirretyn energian arvoa, verkostopituutta sekä asiakasmäärää. Funktiomuodoksi suositellaan lineaarista funktiomuotoa, joka estimoidaan normeerattuun dataa perustuen. Riippuen DEA-menetelmässä tehdystä skaalatuotto-oletuksesta malli voidaan estimoida joko perusmuodossaan nousevien skaalatuottojen oletuksen mukaisena (sisältäen vakiotermin) tai vakioskaalatuottoisena (ilman vakiotermiä).

Perusmallivaihtoehto voidaan siis esittää seuraavasti

kokonaiskustannus = 132 + 0,26 energia-arvo + 0,66 verkkopituus + 0,06 asiakasmäärä

Kokonaiskustannus tuhansissa euroissa riippuu mallin mukaan siis tuotostekijöistä siten, että kaikilla yhtiöillä on yhteinen 132 tuhannen euron peruskustannus, jonka lisäksi jokainen lisäeuro energia-arvossa nostaa kuluja 26 sentillä, jokainen verkkokilometri 660 eurolla ja jokainen asiakas 60 eurolla.

Tässä perusmallissa ei skaalatuottoa koskien ole tehty ennakko-oletuksia. Positiivinen peruskustannus vastaa kasvavia skaalatuottoja eli pieniä yhtiötä koskevaa skaalahaittaa. Malli vastaa siten nousevien skaalatuottojen (non-decreasing returns to scale, NDRS) DEA-mallia.

Mikäli käytetään vakioskaalatuottoista (constant returns to scale, CRS) DEA-mallia, SFA malli voidaan estimoida ilman vakiotermiä, jolloin skaalatuotto-oletukset ovat vastaavat.

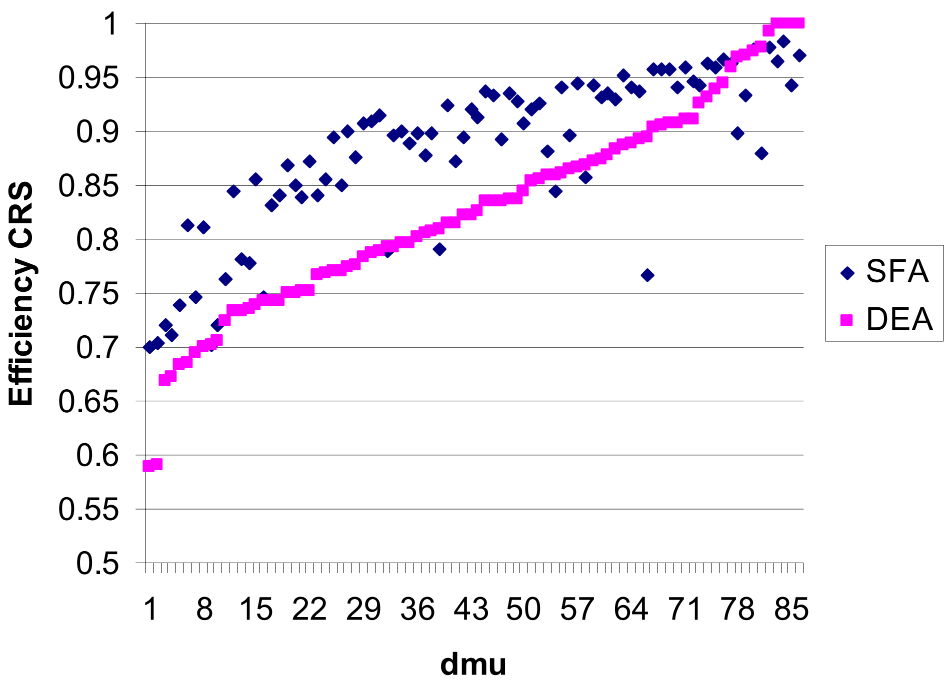
Tässä tapauksessa tuloksena on seuraava malli.

kokonaiskustannus = 0,27 energia-arvo + 0,78 verkkopituus + 0,06 asiakasmäärä

Edellä esitettyyn malliin verrattuna lähinnä verkkopituuden kerroin muuttuu poistettaessa mallista peruskustannuksen määräävä vakiotermi. Nyt yhden verkostokilometrin kustannusvaikutus on 780 €/km.

Vaikka vakiotermi on tilastollisesti merkitsevä perusmallissa, keskeinen johtopäätös on se, että on käytettävä sitä SFA-mallia, joka vastaa vertailukohtana olevaa DEA-mallia. Vakioskaalatuottoisen DEA-mallin kanssa on siten käytettävä jälkimmäistä vakioskaalatuotto-oletukseen perustuvaa SFA-mallia ja vastaavasti kasvavien skaalatuottojen DEA-mallin kanssa perusmallia.

Tarkasteltaessa suositellun mallin tuottamia tehokkuuslukuja, havaitaan, että SFA-mallin tuottamat tehokkuusluvut ovat useimpien yhtiöiden tapauksessa korkeampia kuin vastaavat DEA-luvut. Kuvassa 2 on esitetty samoihin tuotos- ja panostekijöihin perustuvien vakioskaalatuottoisen DEA- ja SFA-mallin tulokset järjestettynä DEA-mallin tuottamien tehokkuuslukujen mukaiseen järjestykseen.




**Kuva 2** SFA- ja DEA-tehokkuuslukujen vertailu, vakioskaalatuotto

SFA-mallin stabiiliuden tarkastelu tehtiin aineiston rajoitteiden vuoksi vuosien 2003 ja 2004 välillä käyttäen mallia, jossa panoksena olivat suositellusta mallista poiketen vain operatiiviset kulut. Tulokset osoittivat, että SFA tarjoaa selvästi vakaamman pohjan vertailulle kuin DEA. Tehokkuusluvuissa tapahtuneet muutokset selittyivät suurimmaksi osaksi yhtiöiden omissa lähtötiedoissa tapahtuneilla muutoksilla. Rintaman muutokset vuosien välillä olivat pieniä.

## Suositellun mallin tasapuolisuus

Suositellun SFA-mallin toimivuutta tarkasteltiin useiden tekijöiden suhteen. Kahden malliversion tuottamia tehokkuuslukuja verrattiin seuraaviin tekijöihin:

- Kokonaiskustannukset, joita käytettiin mallin panostekijänä, koon indikaattorina
- Keskijänniteverkon (6-70 kV) kaapelointiaste erottelemassa taajama- ja maaseutuyhtiöitä
- Operatiivisten kustannusten ja poistojen suhde kuvaamassa verkon arvoa ja operatiivisia kuluja kuvaavia erialisia strategioita
- Asiakaskohtaisia keskeytyskustannuksia kuvaamaan keskeytysten taustalla olevia erilaisia olosuhteista.

Tuloksista voidaan esittää seuraavat keskeiset johtopäätökset.

- Perusmalli suosii lievästi pieniä yhtiöitä ja vastaavasti vakioskaalatuottomalli suuria yhtiöitä. Tekijöiden välinen riippuvuus ei kuitenkaan ole tilastollisesti merkitsevä.
- Kaapelointiaste selittää tehokkuuslukuja tilastollisesti merkitsevällä tasolla. Molemmat malliversiot suosivat lievästi matalan kaapelointiasteen yhtiöitä. Tehokkuus putoaa kesimäärin yhden prosenttiyksikön kaapelointiasteen lisääntyessä 10 prosenttiyksikköä.
- Tasapuolisuuden näkökulmasta sekä perusmalli että vakioskaalatuottomalli ovat yhtä toimivia.

Tasapuolisuuden näkökulmasta malli näyttää toimivan tyydyttävästi. Kuitenkin yksinkertainen funktiomuoto tuottaa yksikertaisemman kuvauksen tuotosten ja panoksen riippuvuudesta kuin DEA-menetelmä samoilla tuotos- ja panostekijöillä. Jatkossa tulisi tarkastella mallin tarkentamista siten, että se huomioi tarkemmin taajama- ja kaupunkiyhtiöiden kustannusajurit. Tällä hetkellä käytössä ollut aineisto ei tarjonnut hyviä lähtökohtia asian ratkaisemiseen. Yhtenä ratkaisuna voitaisiin tarkastella mahdollisuutta jakaa verkkopituus kahdeksi tuotostekijäksi – kaapeliverkoksi ja ilmajohtoverkoksi.

## Tulosten soveltaminen

Kun SFA tulosten käytöstä valvonnassa päätetään, useita kysymyksiä täytyy ratkaista. Ensimmäinen näistä on se, sovelletaanko SFA lukuja varsinaisessa valvonnassa ja jos sovelletaan, miten ne yhdistetään DEA-lukuihin. Raportissa on alustavasti tarkasteltu DEA ja SFA tulosten yhdistämistä laskemalla keskiarvo, tai ottamalla korkeampi tai matalampi tehokkuusluvuista. Näistä vaihtoehdoista keskiarvon laskeminen on tulosten kannalta paras vaihtoehto. Eri lähestymistapoja tulisi kuitenkin tarkastella perusteellisesti ennen päätöksen tekoa.

Asetettaessa tehostamistavoitetta DEA- tai SFA-tulosten perusteella keskeinen kysymys on se, miten tehokkuusluku muutetaan tehostamistavoitteeksi. Tämän hankkeen tavoitteena ei ole ollut esittää mallia tehostamistavoitteen asettamiseksi. Sen asettamisessa on kuitenkin huomioitava operatiivisten ja pääomakustannusten muuttamiseen liittyvät erilaiset aikajänteet. Haasteena tässä on se, että tehokkuuden arvioinnissa panostekijät on yhdistetty ja malli tuottaa yhden tehokkuusluvun. Lähtökohtana tehokkuustavoitteiden asettamisessa voitaisiin käyttää sitä, että arvioidaan eri kustannuskomponentteihin liittyvän tehottomuuden poistamiseen kuluva aika ja lasketaan tämän perusteella vuosittain poistettavissa olevan tehottomuuden osuus. Koska yhtiöiden välillä on eroja operatiivisten ja pääomakustannusten suhteessa, operatiivisten kustannusten ja poistojen välistä suhdetta voitaisiin käyttää indikaattorina eri kustannuserien painottamisessa. Kuitenkin myös tässä tapauksessa eri lähestymistapoja on analysoitava perusteellisesti ennen mallin käyttöä hinnoittelun valvonnassa. Tehostamistavoitteen asettamisessa neljän vuoden valvontajaksolle on myös tarkasteltava mahdollisuutta käyttää useamman vuoden lähtötietoja tasaamaan tietojen vuosittaista vaihtelua.

SFAn ja DEAn tuottamia tuloksia verrattaessa havaitaan, että keskeisenä erona on edellä esitetyn funktiomuodossa esitettävän tehokkaan rintaman lisäksi se, että SFA-mallissa panos- ja tuotostekijöillä on suoraviivainen tuotantofunktiosta laskettavissa oleva vaikutus tehokkuuslukuun. On kuitenkin huomattava, että SFA-mallissa etäisyys rintamaan jaetaan tehottomuuteen ja satunnaisvaihteluun ja siten malli tulee periaatteessa estimoida uudestaan aina aineiston muuttuessa. Tarkasteltaessa pieniä muutoksia voidaan kuitenkin olettaa, että satunnaisvaihtelun sisältävän termin $v_i$ muutokset ovat pieniä. Tällöin uusi tehokkuusluku voidaan kustannustason muuttuessa laskea suoraan seuraavasta kaavasta.

$$E_{i,uusi} = C(y_i) \ / \ (x_{i,uusi} - x_i - C(y_i)/E_i),$$

Kaavassa E viittaa yhtiön tehokkuuslukuun, $C(y_i)$ yhtiön i tuotoksilla laskettuun tehokkaaseen kustannukseen ja $x_i$ yhtiön i kustannustasoon. Alaviite uusi viittaa uusin arvoihin muutosten jälkeen. Myös tuotosten marginaalisia muutoksi voidaan tarkastella laskemalla tehokas kustannustaso kaavan osoittajassa uusien tuotosten perusteella.

Hankkeen tulokset osoittavat, että SFA-menetelmä soveltuu suomalaiseen valvontajärjestelmään käytettäväksi DEA-menetelmän rinnalla. Kuitenkin mallin tarkat parametrit ja yhtiökohtaiset tehokkuusluvut tulee arvioida uudelleen viimeisimmän aineiston perustella ennen mallin käyttämistä varsinaisessa valvonnassa. Samalla on huomioitava aineiston määritelmissä tapahtuneet muutokset kuten keskeytysaikojen muuttunut tilastointi. Myös mallin soveltamiseen liittyvät vaihtoehdot tulee analysoida perusteellisesti ennen mallin mahdollista käyttöä valvontajärjestelmän osana.

# 1 Introduction

The Energy Market Authority (EMV) supervises and promotes functioning of the electricity markets in Finland. As a part of this task it is responsible for supervising the terms and prices of electricity network services. The amendment of the Electricity Market Act in 2004 caused significant changes in the supervision system. In the new regulatory system introduced for the first regulatory period in 2005-2007, the reasonableness of the pricing is supervised on the basis of variety of methods that aim at defining reasonable capital and operational costs of the distribution system operators (DSOs).

EMV has launched a development project where it aims at revising the new regulatory system for the next period 2008-2011 and getting better understanding of those factors that affect distribution business but cannot be taken into account in the monitoring of the reasonableness of the pricing. The project includes, among others, tasks for developing the efficiency measurement model for electricity distribution, presenting a plan for assessing the long-term efficiency improvement potential of the DSOs, introducing company specific efficiency improvement targets, and assessing the incentive impact of various models and procedures.

As a part of this development project EMV has commissioned studies that aim at developing the efficiency measurement of DSOs. Study A aims at further developing the efficiency measurement model based on the DEA method, and study B aims at developing an analogous alternative efficiency measurement model. More specifically study B aims at defining a second efficiency measurement model that can be used for controlling the results of the DEA model. The purpose is to develop a model that EMV can use independently for verifying the efficiencies and efficiency improvement potentials of individual companies. The method and specification will be selected and motivated during the study.

This report is the final report of study B. The report includes a review of the Finnish regulatory system where the developed model is aimed to be included, presents a literature review of alternative efficiency analysis approaches, especially Stochastic Frontier Analysis (SFA) and discusses the ways the alternative model can be used. The key finding in this part is that Stochastic Frontier Analysis (SFA) is the most suitable alternative to Data Envelopment Analysis (DEA). Furthermore, the report describes the analysis process that was used for building the analogous efficiency analysis model based on SFA, presents the conclusions from the analysis of various alternative models and gives a recommendation on the model variant to be used. The regulatory implications related to the recommended model are also analysed.

The report is structured as follows. Section 2 presents the results of the literature review. Section 2.1 discusses the current Finnish regulatory system. Section 2.2 discusses possible efficiency evaluation approaches, and section 2.3 provides a model detailed introduction to SFA. Section 2.4 shortly discusses the various ways of using the alternative model based on the Norwegian and Austrian experiences. Section 3 provides the main results related to choosing the model specification to be used in the analogous model. The regulatory implications of the recommended model are analysed in Section 4. Finally, section 5 presents a short summary of the conclusions and recommendations.

# 2 Review of the regulatory framework and efficiency analysis methods

## 2.1 Finnish regulatory framework

### 2.1.1 Introduction

The reform and deregulation of the Finnish electricity market started in 1995. After the new Electricity Market Act (386/1995) entered into force, the deregulation of the electricity market has taken place in stages. Production, sales and foreign trade have been opened for competition, and transmission and distribution defined as natural monopolies. Finland has joined the Nordic electricity market; networks have been opened to all the customers; production, retail sales and distribution were unbundled; and transmission was centralized to one company.

This study concentrates on the regulation of local distribution system operators (DSOs) that are responsible for electricity distribution at the local level. The networks mainly consist of 20 and 0.4 kV lines. Some distribution companies have 110 kV lines and power stations, and also other voltage levels are used in some cases.

Along with the structural development, the number of DSOs has decreased drastically from the original 200 companies in the past 20 years. Currently, there are 91 distribution network operators (EMV, 2005). Fortum and Vattenfall are the biggest distribution companies.

The Energy Market Authority[4] (EMV) was established to supervise the functioning of the deregulated electricity market and electricity network operations. EMV is a subordinate to the Ministry of Trade and Industry.

### 2.1.2 Legislative framework and key institutions

The primary purpose of the Energy Market Act (386/1995) is to ensure preconditions for an efficiently functioning electricity market and to secure sufficient supply of high-standard electricity at reasonable prices. The primary means for this are to secure a sound and well-functioning economic competition in electricity production and sales, and reasonable and equitable service principles in the operation of electricity networks. The act entered into force in 1995, and minor amendments were made during the years (1018/1995, 332/1998, 138/1999, 466/1999, 623/1999, 444/2003, 1130/2003). In 2004, the act went through a major reform (1172/2004).

The act (9 - 10 §) obliges the distribution companies to

1) Transmit the energy that the customers in the area need against a reasonable compensation
2) Connect all the customers – consumers and producers – to the network against a reasonable compensation
3) Develop the network in accordance with customer needs and so that sufficiently high quality and reliability are achieved.

---

[4] until year 2000 the Electricity Market Authority

The main focus in the development of the regulation system has been the reasonableness of the pricing. In the 2004 reform, the supervision of the distribution prices moved from yearly case-by-case ex-post approach to partly ex-ante supervision that covers all the companies and is based on regulatory periods of 3-4 years. One of the main purposes was to meet the EU requirement of defining the methodology of the regulation in advance and processing times. In the future, the companies are obliged to return to the customers any excess profit for the completed regulatory period through pricing in the next regulatory period. Also the rules for unbundling were tightened and the Market Court was introduced as the first step in the appeal process.

The Electricity Market Decree (518/1995, and amendments 451/1997, 438/1998, 182/2004, 1174/2004) includes more specific rules and regulations on the electricity network licenses and responsibilities of the license holders, the construction of networks, retail sale of electricity, and balance responsibility and balance determination. Also the decree was refined in 2004.

In addition to the Energy Market Act and Decree, the Ministry of Trade and Industry has given specific ministerial decrees and decisions on the unbundling of electricity business activities, instruction on reporting obligations, use of load profile system, invoicing, terms of connecting etc.

EMV has the responsibility of implementing the rules and regulations set by the legislation and the ministry. The authority has the responsibility for making many decisions related to the actual implementation of the regulation system and make decisions whether the companies operate according to the set rules.

The key trend since the deregulation of the market has been gradual tightening of supervision. New issues have been added to the regulation system and regulation has become more detailed. The next sub section discusses the actual changes.

### 2.1.3 Regulatory system

The current regulatory model is based on the guidelines for reforming supervision of electricity and gas network operations that were set out in a report of the Ministry of Trade and Industry working group for reform of regulation of pricing in the energy market. The actual system was developed by EMV and it is introduced in a separate document that forms the basis for the methodology decisions given prior the start of the regulatory period. (EMV, 2004b)

The current system is based on an ex-post rate of return regulation. In addition, the operational costs must be reasonable. This interpretation is based on the Electricity Market Act and its preambles.

The first and still applicable part of the regulation is the rate of return regulation. Initially, the rate of return was supervised based on one year periods. The reasonable rate was based on adjusted financial statements and a Weighted Average Cost of Capital (WACC) model. The first decision on the return level was made in 1999 (case Megavoima Oy) and the Supreme Administrative Court confirmed this decision, and hence the used approach and the power of EMV, in 2000.

After the implementation of the rate of return regulation, EMV developed a model for ex-post yardstick regulation of operational costs. The levels of reasonable operational costs were defined on the basis of a DEA model (Korhonen et al. 2000). Excess costs were interpreted as profit in the early ex-post evaluation of the rate of return. Eventually, this model was used only in cases where the network operators benefited from it, and the system for supervising operational costs was completely reformed in 2004.

In the last reform in 2004 many components of the regulation were changed. The rate of return is regulated based on the same basic principles but many details like rules for depreciation changed. Regulatory periods were introduced and the first period is 2005-2007. A new approach to the regulation of costs was introduced, and now the reasonable cost level is based on a cost cap. The current model also includes an obligation to return the excess profit to the customers during the following period instead of just changing the tariffs after the supervision decision. On the other hand, the system allows higher return during the next regulatory period if the return has been below the reasonable earnings level. The reasonable earnings level will be calculated on the basis of amount of capital, reasonable rate of return, and adjusted profit and loss account.

In the beginning of the first regulatory period, the amount of capital is defined based on the net present value (NPV) of the network. This is calculated by multiplying the replacement value of the components by the ratio of the average age and the holding time (i.e. straight-line depreciation). The replacement value is dependent on the type of component and the environment (urban, semi-urban or rural), and the holding time of the component groups can be chosen within certain limits. For the two following years, the net present value is adjusted based on straight line depreciation and actual investments (valued with standard prices). Other assets related to network business are valued at book value, and financial assets are excluded.

The reasonable rate of return is based on a Weighted Average Cost of Capital (WACC) model. The reasonable rate of return on equity is calculated on the basis of a Capital Asset Pricing (CAP) model, i.e. reasonable rate is risk free rate + levered beta factor multiplied by a market risk premium. In the implemented model, the risk free rate corresponds to the 5-year Finnish government bond. The levered beta is 0.395 or 0.429 depending on the ownership (i.e. taxes) of the company, and the market risk premium is 5%. The reasonable rate for dept is risk free + 0.6%. The capital structure is assumed to correspond to dept/asset ratio 30/70 for all the companies. The reasonable rate is calculated separately for each year. For example for year 2005, it is based on government bond May 2004 average, which is 3.53%. When the WACC model is applied to the rates and capital structure above, the reasonable rate of return on the total assets in 2005 is 4.77% or 5.21% depending on the ownership.

The acceptable costs are based on a cost cap, which is defined ex-ante based on the historical operational costs of the company and a CPI-X factor. The X-factor is based on industry level productivity development (frontier shift) and was defined with a DEA based Malmquist analysis. The X factor is 2.2% and the price index (actually industrial production price index) has changed 0.9% p.a. on average in 1995-2002. The reference cost levels in the CPI-X model are the average operational costs in 2000-2003. If the volume of the operations has changed, the cost level is corrected on the basis of the change in the network volume and number of customers. The current model does not include a company specific component in the X-factor. This is one of the important development areas.

During the regulatory period, EMV makes yearly calculations for all the companies, but the official supervision decisions will be made only after the end of the period. Hence the actual decisions that are based on rules described above are made ex-post. Customers have no real role in the regulation of rate of return and costs, as all the companies are automatically supervised by EMV. In the new system, the case by case discretion has decreased. However, there is still some flexibility in defining the asset base and this is a potential source of conflict.

In the decision, EMV can oblige the company to change its tariffs so that these are reasonable and the windfall profits from the previous period will be compensated for. The company will make the actual decision concerning the tariff levels and tariff structure.

### 2.1.4 Development issues

In 2005 EMV launched a development project to further develop the current regulation system. As a part of the project three external sub-projects related to efficiency measurement of DSOs have been commissioned. Study A aims at further developing the efficiency measurement model based on the DEA method. This study concentrated on four main aspects: 1) how technical quality of the distribution will be included in the DEA based efficiency evaluation and the regulation system in general; 2) how investments and capital costs are included in the efficiency evaluation and regulation system; 3) how the model should be developed so that the long term efficiency improvement potential during the regulatory period can be defined; and 4) how exceptional companies should be treated in the DEA model. Study B aims at developing an analogous alternative efficiency measurement model. The model should provide a comparison point for the new model developed in study A.[5] Finally study C concentrates on exploring the possibilities to take customer service into account in the supervision of the reasonableness of the pricing in the future.

In the next regulatory period, EMV aims at introducing a company specific X-factor that would reflect the efficiency improvement potential of the company. If we look at studies A and B from the regulatory system point of view, one of the key goals is to develop a yardstick approach for defining a company specific X-factor for the next regulatory period. Furthermore, the current regulatory model (described in section 2.1) does not take quality into account. (There is a separate compensation scheme, however). Hence study A includes also a wider perspective of developing the whole regulation system. However, EMV has the main responsibility for deciding on the way the efficiency scores will be transformed into X-factors.

The objective of study B is to develop and define a secondary efficiency measurement model that can be used for controlling the results of the DEA model used by EMV. Decisions on the actual way of using the results in regulation will be made by EMV based on the internal development project the runs in parallel with the efficiency benchmarking projects. The aim is that the secondary model could be used in parallel with the DEA model in the actual regulation. Hence the data and specifications used the DEA model will serve as a comparison point for the alternative model. However, the actual model that will be used in regulation has not been defined yet, and study A will propose changes especially related to capital costs and quality. Hence the results of study A will describe the regulatory context where the alternative efficiency evaluation model will be applied.

The current plan is that the DEA model and the alternative model would be used in setting the company specific component of the X-factor. Currently the X-factor in the model is applied to controllable operational expenditure (opex) and it does not include a company specific component. In the future the X-factor may be applied to controllable opex, capital costs and quality costs caused for customers (interruption costs). The decisions on the actual use of the models will be made by EMV after the projects have been finished.

---

[5] Studies A and B have been implemented in parallel. The results of study A are presented in a report by Honkapuro, Tahvanainen, Viljainen, Lassila, Partanen, Kivikko, Mäkinen, Järventausta (2006)

.

The main concern in study B is the potential estimation error caused by DEA method, and hence the model specifications would ideally be analogical to the DEA model that will be used in regulation. However, the method and specifications should be motivated and validated independently. Identification of problems in the model structure or data is a secondary objective. The basic principle is that the results of study B will be public, but the decisions on the way of applying the results and the possible future use of the alternative model will be decided later by EMV.

## 2.2 Possible efficiency evaluation approaches

### 2.2.1 Overview

At a general level, one can distinguish between parametric and non-parametric models on the one hand and between stochastic and non-stochastic models on the other. We first discuss the different types of benchmarking models and we briefly summarize their pros and cons. The main purpose is to offer a discussion of some of the factors that we consider to be of particular importance in regulatory applications. The vast number of scientific papers on frontier analysis in general and DEA methods in particular[6] prohibits a balanced and comprehensive coverage of benchmarking approaches within any project.

***Parametric versus non-parametric***

In the modern benchmarking literature parametric models are characterized by being defined a priori except for a finite set of unknown parameters that are estimated from data. The parameters may refer to the relative importance of different cost drivers or to the parameters in the possibly random noise and efficiency distributions. The non-parametric models are characterized by being extremely flexible in terms of the production economic properties that they invoked. Only a broad class of functions – or even production sets – are fixed a priori and data is used to estimate one of these. In this case, the classes are so broad as to prohibit a parameterization in terms of a limited number of parameters.

***Deterministic versus stochastic models***

Stochastic models are the most flexible in terms of the assumptions one can make about data quality. One makes a priori allowance for the fact that the individual observation may be affected by random noise, and tries to identify the underlying mean structure stripped from the impact of the random elements. In deterministic, i.e. non-stochastic models, the possible noise is suppressed and any variation in data is considered to contain significant information about the performance of the unit and the shape of the technology. The deterministic approaches therefore presume data of good quality. On the other hand, in terms of the model structure, the non-parametric deterministic approaches are the most flexible ones.

***Taxonomy***

The two dimensions leads to a 2x2 taxonomy of methods as illustrated in Table 2.1 below. A few original key references are included.

[6] There are by now more than 1000 scientific papers and numerous text books focusing on frontier models, c.f. the bibliography on www.deazone.com.

**Table 2.1** Model taxonomy

| | Deterministic | Stochastic |
|---|---|---|
| **Parametric** | **Corrected Ordinary Least Square (COLS)**<br><br>Greene (1997), Lovell (1993), Aigner and Chu (1968) | **Stochastic Frontier Analysis (SFA)**<br><br>Aigner, Lovel and Schmidt (1977), Batesee and Coelli (1992), Coelli, Rao and Battesee (1998) |
| **Non-Parametric** | **Data Envelopment Analysis (DEA)**<br><br>Charnes, Cooper and Rhodes (1978), Deprins, Simar and Tulkens (1984) | **Stochastic Data Envelopment Analysis (SDEA)**<br><br>Land, Lovell and Thore (1993), Olesen and Petersen (1995), Weyman-Jones (2001) |

We emphasize that for each class of model, there exist a large set of model variants corresponding to different assumptions about the production technology, the distribution of the noise terms etc. We will discuss some of the key assumption below. We presume a basic knowledge of these models here and do not explain them in detail. We simply recall the differences in a simple cost modelling context. The setting then is that we seek to model the costs that results when best practice is used to produce one or more outputs.

Figure 2-1 illustrates the approaches. Corrected ordinary least square (COLS) corresponds to estimating an ordinary regression model and then making a parallel shift to make all units be above the minimal cost line. Stochastic Frontier Analysis (SFA) on the other hand recognizes that some of the variation will be noise and only shift the line – in case of a linear mean structure – part of the way towards the COLS line. Data Envelopment Analysis (DEA) estimates the technology using the so-called minimal extrapolation principle. It finds the smallest production set (i.e. the set over the cost curve) containing data and satisfying a minimum of production economic regularities. Assuming free disposability and convexity, we get the DEA model illustrated in Figure 2-1. Like COLS, it is located below all cost-output points, but the functional form is more flexible and the model therefore adapts closer to the data. Finally, Stochastic DEA (SDEA) combines the flexible structure with a realization, that some of the variations may be noisy and only requires most of the points to be enveloped.

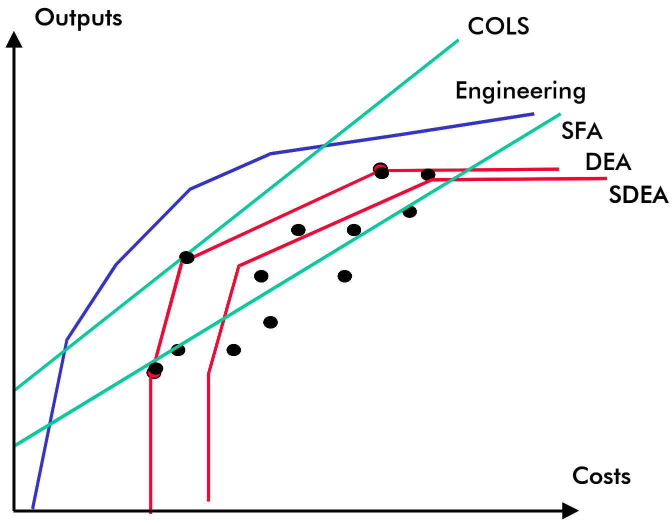



**Figure 2.1** Benchmarking methods (example)

In Figure 2.1 we have included a fifth frontier, termed engineering. The idea is to base the modelling on data from engineers about best possible performance, perhaps in idealized settings. This is often applicable for example when modelling an industrial production process. An advantage of this approach in regulation is that the benchmark norm cannot be affected by the evaluated unit. Also, if properly developed, it may possibly reflect in more details the operating environments of the different DSOs. A drawback of the approach is that it is very expensive to develop. In the case of electricity distribution a key challenge would be to understand and model the non-technical processes. Moreover, and probably more importantly, the historical conditions are difficult to include such that DSOs will tend to be evaluated as if past decisions were taken based on present information about technology, distribution of load and generation etc.

### *Advantages and disadvantages*

We will now focus on the advantages and disadvantages of these methods in general, and in particular their relative merits in a regulatory context.

Some of the strengths of non-parametric and deterministic methods like DEA include

- Requires no or little preference, price or priority information
- Requires no or little technological information
- Makes weak a priori assumptions
- Handles multiple inputs and multiple outputs
- Provides reel peers
- Identifies best practice
- Cautious or conservative evaluations (minimal extrapolation)
- Supports learning and in some cases planning and motivation
- Game theoretical foundation of the industry-regulator relation

Some of the strengths of parametric and stochastic methods like SFA are

- Strong theory of significance testing (sensitivity, re-sampling, bootstrapping, asymptotic theory)
- Separates noise and efficiency
- May leave lower rents when functional form known
- Creates anonymous peers, may be relevant in regulation

As indicated, the different approaches have different advantages and disadvantages. From a regulator's viewpoint, the relative importance of these merits depends on the overall regulatory approach (cf. Agrell and Bogetoft, 2003a), i.e., the role assigned to the model among the regulatory instruments.

In the Finnish case EMV has identified the following weaknesses in the DEA model. Specifically the potential estimation errors caused by noise in the data have been identified when the year-to-year fluctuations of the efficiency scores have been analysed. Furthermore some problems related to exceptional units have been identified. Finally the argumentation behind the choice of inputs, outputs and environmental factors would benefit from more formal analysis.

In our view, a fundamental difference from a general methodological perspective and from regulatory viewpoint is the relative importance of flexibility in the mean structure vs. precision in the noise separation. This means that there are basically two risks for error that cannot be overcome simultaneously. These are 1) risk of specification error, and 2) risk of data error.

Specification error is related to the inability of the model to reflect and respect the real characteristics of the industry. Avoiding the risk of specification error requires a flexible model in the wide sense. This means that the shape of the model (or its mean structure to use statistical terms) is able to adapt to data instead of relying excessively on arbitrary assumptions. The non-parametric models are by nature superior in terms of flexibility.

Data error means inability to cope with noisy data. A robust estimation method gives results that are not too sensitive to random variations in data. This is particularly important in individual benchmarking – and probably less important in industry wide motivation and coordination studies. The stochastic models are particularly useful in this respect.

It is worthwhile to observe that the two properties may to some extent substitute each other. That is, the flexible structure allowed by non-parametric deterministic approaches like DEA may compensate for the fact that DEA does not allow for noise and therefore assigns any deviation from the estimated functional relationship to the inefficiency terms. Likewise, the explicit inclusion of noise or unexplained variation in the data in SFA may to some extent compensate for the fact that the structural relationships are fixed a priori, i.e. the noise terms may not only be interpreted as a data problem but also as a problem in picking the right structural relationship. As an illustration of this we have found in other studies that the SFA efficiencies are often larger than the DEA efficiencies as long as the model is somewhat ill-specified., i.e. the inputs and outputs are badly chosen. The reason is that SFA in this case assigns the variations to the noise term while DEA assigns everything to the efficiency term. As the model is extended to include more relevant inputs and outputs, the two methods have been found to produce quite comparable results.

As both advanced non-parametric approaches like stochastic DEA and parametric approaches like SFA can be used to solve the practical challenges related to the Finnish regulatory context, these two are discussed in more detail below.

### 2.2.2 Non-parametric approaches

Recent developments in the DEA literature do ease the "statistical" evaluation of DEA results. This means that the draw-backs of the DEA method indicated above, viz. significance testing, sensitivity to noise, to bias in the parts of the input-output space where the number of observations are few etc can now be at least partly overcome.

To investigate the sensitivity of the DEA model to (changes in) the industry structure, one can examine how the efficiency of each unit depends on the other units available. This can be done in a systematic way using so-called efficiency step ladders (ESL), see Edvardsen (2004). The idea is to investigate the impact for each unit of eliminating one peer unit at a time.

DEA models provide cautious estimates of the saving potentials and cost inefficiencies. This is one of the attractive features of DEA and is part of the theoretical foundation for the optimality of DEA based yardstick competition, cf. Bogetoft (1997, 2000) and Agrell, Bogetoft and Tind (2004). If the model structure and the variables are chosen correctly, it means that no-one will be required to produce at costs that are below the truly minimal ones. The back-side of the cautiousness is that the cost efficiency is biased upwards. On average, the units will look more efficient than they really are. Moreover, the bias for some units may be larger than for others. This means that some distribution companies may be facing tougher norms than others. This is an unattractive feature also in regulation since it may prohibit a fair treatment of the different units.

Before getting too concerned it is important to understand the implications. All cost norms should suffice by the cautiousness property, but for some units the extra benefits from a biased evaluation is more lucrative than for others. Hence, in summary, all units benefit from the doubt in the DEA model – but the benefit is not equally distributed among the firms.

The recent advances in the statistical foundation of DEA allow us to estimate and hereby possibly control for the bias. The expected difference between the DEA frontier and the true frontier, the efficiency bias (BiasEff) can be calculated by boot-strapping. Combining this with the observed efficiencies we get a Bias corrected efficiency (CorrEff), i.e. the efficiency in relation to an estimate of the true frontier.

Bootstrapping also allows us to estimate the standard deviations of the bias corrected DEA efficiency estimates. Intuitively, one would expect that the standard deviation of the bias corrected efficiency estimates are larger, the more specialized the unit (i.e. the less densely populated is the part of the production space where it operates). Instead of using standard deviations, one can also and more correctly use the boot-strapped distributions to pick out the confidence intervals without presuming normal distributions.

Stochastic DEA, SDEA, is another way to overcome some of the drawback of standard DEA models. The idea of SDEA is that data is noisy and that this should be reflected in the construction of benchmarks and costs norms. In a SDEA approach, one would – without introducing new a priori assumptions about the underlying "mean" cost structure find the saving potential that we are e.g. 95% certain is possible. Although this sounds promising by combining the advantages of both the stochastic and the non-parametric approaches, the theory of SDEA is still insufficiently developed. The linkage with regulation is not developed, and the technical assumptions about noise distributions needed to make practical implementation are very restrictive. It is therefore our evaluation, that SDEA is an interesting approach, not the least academically, but that it is not yet ready for practical usage in a regulatory context.

### 2.2.3 Parametric approaches

In the parametric SFA approach, the separation of noise and inefficiency is technically done by assuming that the noise is two sided and inefficiency is one sided. Inefficiency makes costs increase and makes production fall short of the best possible, while noise may also lower the observed costs or increase the observed output. In addition to having one- and two sided deviations, the separation of noise and inefficiency is accomplished by making specific assumptions about the nature of the distributions, e.g. normal and half normal.

In the parametric approach, one also makes specific assumptions about the type of relationship between the inputs and outputs. The so-called functional form may for example be linear, log-linear or translog. We shall return to these assumptions below.

To be more specific, we may distinguish between three combinations of noise and inefficiency. Namely pure noise models, pure efficiency models and combined models. In a cost setting, we may assume that costs, x, depend on a series of output driver, y, as well as on a combination of the inefficiency term $u \geq 0$ and the noise term v for each of the DSOs i

- Pure noise (Ordinary least squares (OLS), average cost function): $x_i=C(y_i) + v_i$
- Pure inefficiency (Deterministic frontier): $x_i=C(y_i) + u_i$
- Combine (Stochastic frontiers): $x_i=C(y_i)+ u_i + v_i$

In the specifications above, C(y) is the minimal costs function. It defines the least expensive way to provide the outputs y. The functional form of C(y) is given, except for a some unknown parameter values β, i.e. one uses C(y,β). The statistical analysis seeks to estimate the functional relationship, i.e. β, and to estimate the inefficiencies, i.e. $u_i$..

The first of these specifications (OLS) is the specification in classical statistics. It fits a function to the data in such a way that the positive and negative deviations are as small as possible. The standard measure of goodness-of-fit is the sum of squares of deviations, which is why this approach is often referred to simply as the OLS, ordinary least squares approach. Since the OLS approach does not work with the idea of individual inefficiencies, the usage of OLS in regulation is problematic. It can of course be used to identify likely cost drivers and to evaluate structural inefficiencies. Individual inefficiencies, however, are by assumption absent. In many cases, therefore, OLS estimates are only a starting point. It is followed by some ad hoc adjustment of the OLS estimate towards the frontier. Conceptually and theoretically, it would be better to do so as part of an integrated approach as the SFA approach.

The deterministic frontier approach assumes that there is no noise, only inefficiency to explain deviations from the model costs C(y). The functional form is furthermore specified, e.g. to be linear. DEA has the same starting point in terms of noise, but works with a very large, in fact non-parametric, class of functions to begin with. In the case of the DEA variable returns to scale model, for example, the only a priori assumption is that C(y) is weakly increasing and convex. No more specific assumptions are made about the functional form.

One can therefore say that deterministic parametric models have the disadvantages of DEA, no noise, without the advantages of DEA, namely a flexible functional form. We shall therefore not expand too much on these approaches.

The third approach, stochastic frontier approach allows for both noise and inefficiency. The advantage of the approach is first of all that the idea is nice. Conceptually, it is attractive to allows for the realistic existence of both noise and inefficiency. The specification is also attractive by allowing the use of classical statistical approaches like maximum likelihood estimation, likelihood ratio testing etc.

The drawback of the approach is on the other hand that we need a priori to justify 1) the distribution of the inefficiency terms and 2) the functional form of the frontier.

In the next section, we will expand on the SFA approach, and we will discuss how to cope with the drawbacks of having to specify a functional form and inefficiency distribution a priori.

Before turning to the details, one general observation is worthwhile. It is often believed that acknowledging noise means that the evaluated units are put in a more favourable light. The double sided nature of noise, however, implies that a DSO may come worse out of the evaluation using a model with noise than they do in a model without noise. The data may suggest that the high performance is in part the result of good luck (cost decreasing noise for example) such that the real performance is actually worse. In game terms, the evaluated can claim bad luck in case of bad performance – but the regulator can claim good luck in case of good performance. This is especially relevant in the case of extreme units.

### 2.2.4 Basic illustrations

The ideas of these parametric approaches are illustrated in the figures below.

In Figure 2.2, the classical statistical approach is illustrated. The cost function models the average or expected costs for different activity levels and any deviation is attributed to random errors or noise (i.e. all the DSOs are supposed to be equally efficient). The high costs of DSO 1 are therefore the result of bad luck, an idiosyncratic random extra cost of $v_1$ while the low costs of DSO 2 is the result of good luck corresponding to a random saving of $v_2$. Of course, if we believe in the conditions under which we estimate the cost function, we also have no basis for any regulatory interference since every deviation from the cost norm is considered to be beyond the control of the DSOs.

In Figure 2.3, the deterministic approach, we take the opposite attitude and associate any deviations from the cost norm to inefficiency. Since inefficiency will always increase the costs, the cost function is now a classical frontier function, located below all observations in the cost function interpretation. Thus, the high costs of DSO 1 is believed to be the result of a high DSO specific inefficiency $u_1$ while the relatively small marginal extra costs of DSO 2 indicates that the inefficiency of DSO 2, $u_2$, is small.

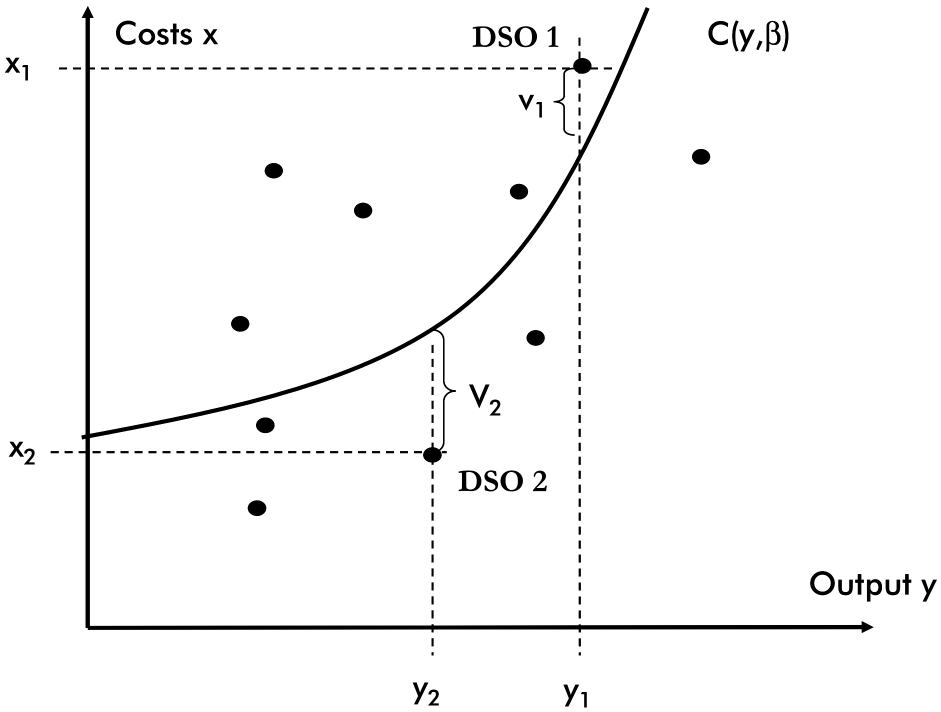



**Figure 2.2** Average cost (all noise)

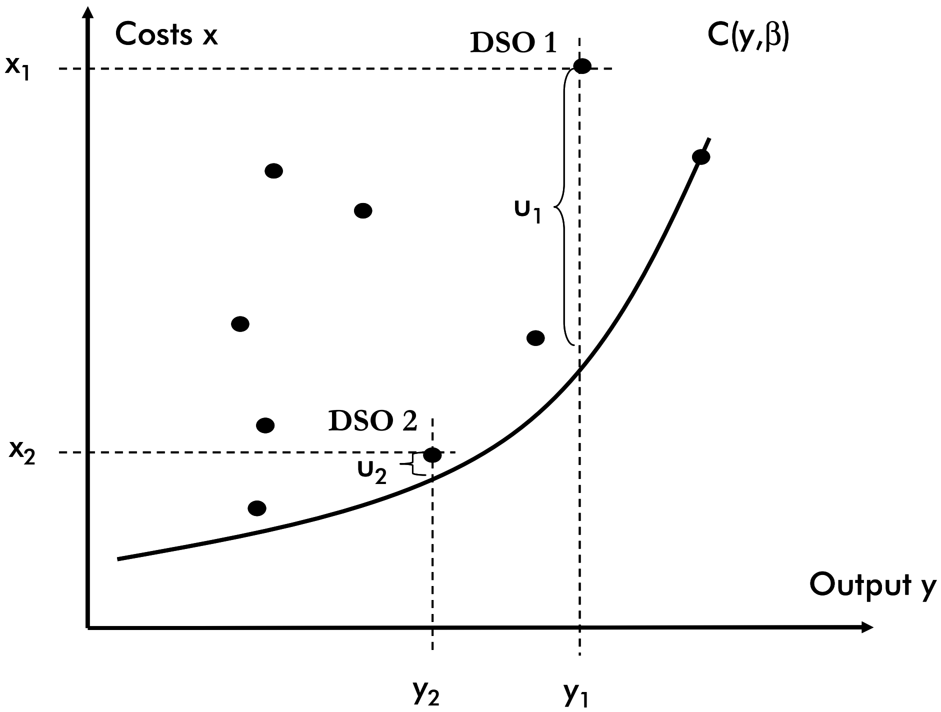



**Figure 2.3** Deterministic frontier (all inefficiency u)

In Figure 2.4, the stochastic frontier approach is illustrated. The minimal cost of producing given output levels are again depicted by the cost function $C(y,\beta)$, but now deviations from the minimal costs is generated by two factors, the noise represented by the symbol v and the inefficiency represented by the symbol u. We see that the excessive costs of DSO 1 are generated by both positive noise (bad luck) and by inefficiency. The impressing performance of DSO 2 on the other hand is estimated to be the result of very good luck leading to a cost reduction of $v_2$ combined with some inefficiency leading to a slight increase of $u_2$ above the good luck outcome.

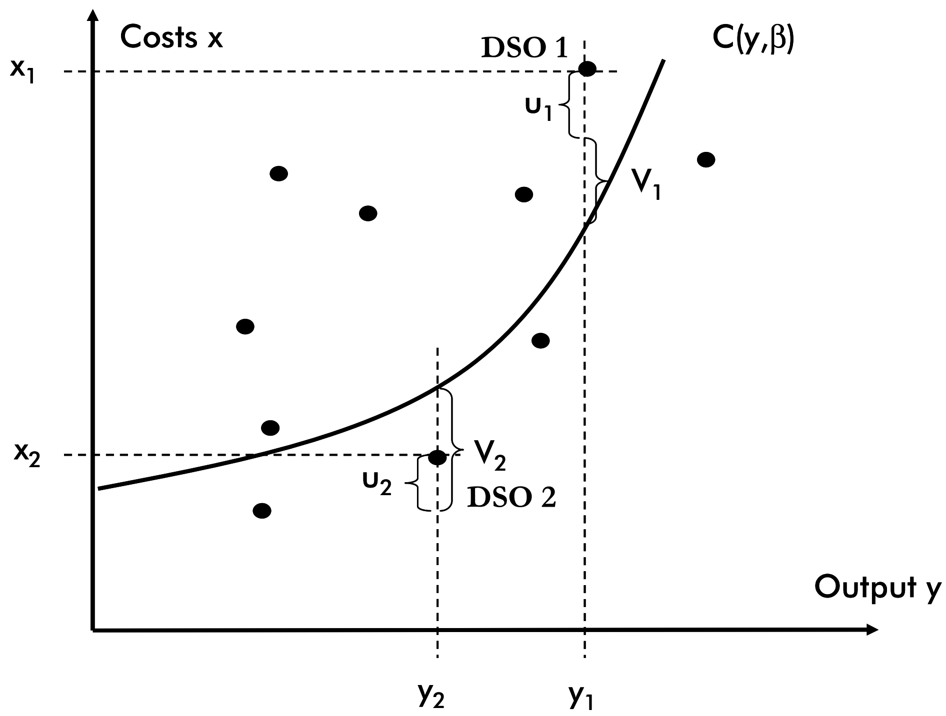



**Figure 2.4** Stochastic frontier (both noise v and inefficiency u)

The determination of the cost function or more precisely the cost function parameter β as well as the noise and inefficiencies, v's and u's, is in general done using so-called maximum likelihood estimation. This is the fundamental principle of statistical estimation and is based on the simply logic that we believe in the functions and noise and inefficiency values that make the observations we have as likely or probable as possible. That is, if we have two possible functions for example, and the observations we have often result from the former and only seldom from the latter, than we infer that that the first function is probably the correct one. Of course, we never know this with certainty since we can never observe the function nor the noise and inefficiency terms directly, but we have good reasons to believe in the ones with the highest likelihood of generating what we can actually observe, namely the actual outputs and costs levels.

The likelihood principle is not only the basis of classical statistical theory. It is interesting that the same logic also applies in a contracting or regulation context. It can be shown quite generally that the optimal regulation should be based on maximum likelihood estimates, cf. Holmstrom (1979, 1981) and with direct linkage to benchmarking, Bogetoft (1994, 1997).

### 2.2.5 Conclusion

In the choice of an efficiency evaluation model for regulation, there is a fundamental choice between the risk of misspecification of cost structure – how cost depends on a series of cost drivers – and the risk of noise in the data – how reported performance may deviate from actual performance.

The DEA approach is superior in terms of the specification problem while the SFA approach is superior in terms of the noise problem. Ideally, we would like to use flexible models that are robust to random noise. The problem however is that this is rarely possible in practice. The estimation task becomes bigger, the data need larger and still we cannot avoid a series of strong assumptions about the distributions of the noise terms. Hence, coping with uncertainty requires us to dispense somewhat with flexibility and vice versa. Approaches like SDEA that aim at combining flexibility with robustness to random noise are also still theoretically undeveloped and mainly of academic interest. Furthermore, SDEA or the stochastic extensions to DEA do not provide a complete alternative to DEA and hence do not provide an independent comparison point.

Hence our conclusion is that the SFA approach is the obvious choice as a complementary model to the DEA model used on the Finnish regulatory context. In the following, we will expand on the SFA approach. We will describe it in more technical terms and we will outline the principal ways to cope with the limitations of the SFA approach, viz. how to choose a flexible functional form and how to choose a distribution for the inefficiency elements. We will also discuss the role of the SFA approach as a complement to the DEA model and how the dual usage of such models has been considered in Norway and Austria.

## 2.3 Introduction to Stochastic Frontier Analysis

The stochastic frontier approach was introduced independently by Aigner, Lovell and Smith (1977) and Meuser and Van den Broeck (1977). This chapter provides a short introduction to the basic characteristics of SFA. A comprehensive introduction to SFA can be found for example in Coelli et al. (1998).

The key feature of SFA is that it allows for both noise and inefficiency. In a cost function interpretation, it would specify the costs as

$$x_i=C(y_i)+ u_i + v_i$$

where inefficiency $u_i$ is half-normal $N_+(0, \sigma_u^2)$ and noise term $v_i$ is normal $N(0, \sigma_v^2)$.

The advantage of the approach is first of all that the idea is nice. Conceptually, it is attractive to allow for the realistic existence of both noise and inefficiency. As indicated by the specification, the idea is that there is some underlying cost function that gives the minimal costs for different output levels under normal operations. In a DEA analogy, this corresponds to the DEA frontier. This costs level is rarely observed, however. Actual costs deviate from the true minimal costs for two reasons.

One is noise or measurement errors. This terms picks up for example random variations in the number of breakdowns that require more or less maintenance, variations in salary levels, periodization choices, small registration mistakes etc. In short, it picks up all the small effects that are not modelled in full details. These effects can sometimes produce too small and sometimes to large costs levels. The noise term is therefore assumed to be symmetric around 0, i.e. we assume that it may just as well lower as increase the actual costs compared to a situation with normal conditions.

The other term that makes actual costs deviate from minimal costs is inefficiency. This term is intended to pick up the coordination and motivation problems in different firms – and the fact that some firms are doing a better job solving these problems. A stylized interpretation of inefficiency is that it captures the "inability" or "laziness" of managers. It is more correct however to interpret the inefficiency as the extra costs that are the results of the internal organization problems of coordinating activities and of motivating the employees. Part of this can also be a lack of bargaining power towards firms that the company in question outsource to. In case the companies are not operating in a homogeneous environment and the model does not include factors describing the differences in the conditions, inefficiency can partly be caused by the external conditions.

The SFA specification is also attractive by allowing the use of classical statistical approaches like maximum likelihood estimation, likelihood ratio testing etc. In a SFA approach, we use the data to come up with a best estimate of the underlying costs function C. Compared to DEA, we have less freedom in our choices since we have to decide already at the outset about a possible classes of such functions. Given a best estimate of the cost function C, we can determine the noise plus inefficiency by comparing the actual cost and the cost function value. In SFA, we next investigate the distribution of these actual deviation u+v, and we come up with a best splitting of these terms into its parts, the inefficiency u and the noise v. This splitting is done using the classical statistical principle of maximum likelihood estimation once again. In practice, the cost function and the splitting are determined simultaneously using an iterative process. The ability to distinguish noise and inefficiency hinge on the assumption of one being symmetric and the other being asymmetrically distributed.

In comparison to DEA, it is important to understand that DEA does not have any assumptions about the distribution of the inefficiencies across firms – and DEA does not believe that the data contains noise.

There is one more difference between DEA and SFA in terms of inefficiency. A common DEA measure of cost efficiency is

minimal cost / actual cost

So efficiency is a relative term. An efficiency score of 0.5 shows that the company uses 100% above minimal costs.

The u term in the SFA models is different since it is added. It gives the absolute amount of extra costs, say 1.7 million Euro. To make the results comparable with DEA, we shall usually present not the u terms, but the corresponding relative efficiencies. In principle, this is calculated as

$$C(y_i) \,/\, [C(y_i) + u_i]$$

This ratio is analogical to the above ratio for DEA and corresponds to Farrell efficiency score (Farrell distance). An alternative approach would be to use Shepard distance that is defined as actual cost / minimal cost and hence scores above one refer to inefficiency.

The drawback of the approach is on the other hand that we need a priori to justify 1) the distribution of the inefficiency terms and 2) the functional form of the frontier. We will now discuss how to cope with these difficulties.

### ***Different inefficiency distributions***

SFA requires some a priori assumption about the distribution of the inefficiency term in order to separate noise and uncertainty. On the other hand, it is hard to give strong arguments for a specific form like the half normal. It is therefore better to start with a more general and more flexible specification and to let the data reveal as closely as possible the correct distribution.

A good alternative to the original specification above ($u_i$ is half-normal, i.e. $N_+(0, \sigma_u^2)$) is to assume a truncated normal distribution. That is, one can assume that $u_i$ is $N_+(\mu, \sigma_u^2)$, i.e. a normal distribution centered around $\mu$ and next truncated to be above zero. In case $\mu$ is equal to zero these assumptions coincide. However, using data to estimate just one more parameter, the $\mu$ value, allows for quite a flexible starting point. This is illustrated in figure 2.5 below. When $\mu$ (mu in the figure) is negative, the underlying normal distribution is centred to the left of the y-axis and the truncated positive parts are therefore monotonously decreasing as we see for the boldest illustrations below. This means that most of the units would have high efficiency score while only a small part of the group clearly differ from the main group and are very inefficient. Using a positive $\mu$ instead we get inefficiency distributions more like a traditional normal distribution except that the left tails are truncated as for example the dotted distribution illustrates. This means that efficiency score are more evenly distributed along the whole range.

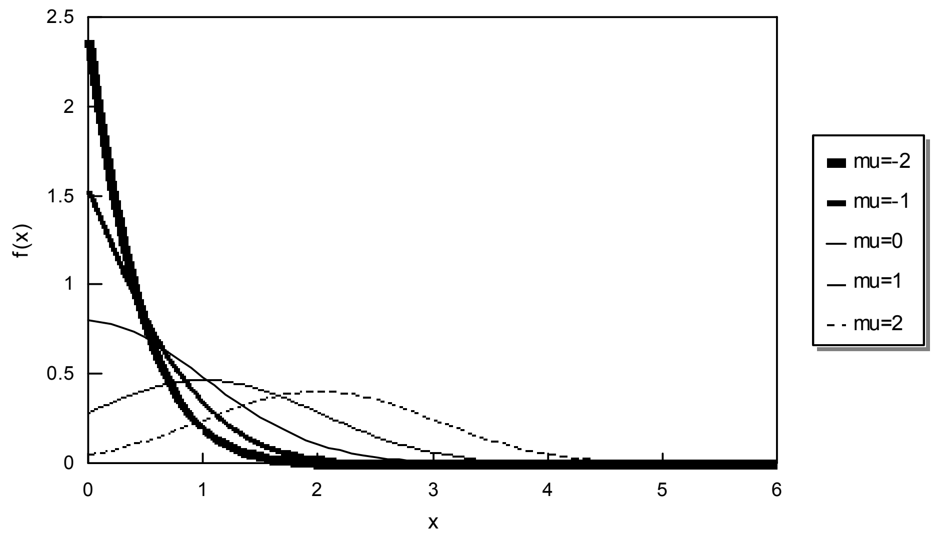



**Figure 2.5** Alternative inefficiency distributions (Bogetoft and Otto, 2005) (mu referes to $\mu$)

The truncated normal distribution is not only relevant because it is a flexible starting point. It is also convenient when there is data from several periods or one wants to account for factors that the DSOs cannot affect but which may affect DSO performance.

Battese and Coelli (1992) proposed a specification of the inefficiency distributions that allows for changes over time in a panel (i.e. a dataset consisting of several time periods). This means that separate $\mu$ value is estimated for each period.

Another variant of the inefficiency distribution is to assume that it depends on a vector of firm-specific variables, including environmental variable. This offers an alternative to the most obvious way of including the environmental variables in the model as output factors. Such suggestions have been developed by Kumbhakar, Ghosh and McGukin (1991) and Reifschneider and Stevenson (1991) among others. A particular simple form is given in Battese and Coelli (1995). They assume that $u_{it}$ is $N_+(z_{it}\delta,\sigma_u^2)$, where $z_{it}$ is a p×1 vector of variables which may influence the efficiency of DSO i and $\delta$ is an 1×p vector of general impact parameters to be estimated. This is a particularly attractive specification since it allows the second stage analysis of explaining variations in efficiency to be integrated in the first stage of estimating the efficiencies.

That is, the somewhat ad hoc nature of the so-called second stage analysis in DEA studies is eliminated by this specification.

***Functional forms of a cost function***

The second fundamental problem in a parametric frontier approach is to select a functional form for the frontier. The selection of functional form is guided by intuition and data as well as theory. An experienced statistician is usually good at choosing functional forms with possibly data transformations and the sufficient degrees of freedom to provide a reasonable goodness of fit of the data at hand. In addition, theory guides the selection by imposing reasonable properties on the estimated function, e.g. that costs function is homogenous in prices or that output sets are convex.

A good general principle is to use the simplest possible representation with the sufficient flexibility to represent data. The simplest possible form is the linear one and a good starting point – and even a starting point used in the iterative procedures used to estimate more advanced forms – is therefore recommended to do a linear regression of cost on the different outputs.

A slightly more complicated specification is the log-linear one being linear in the log of the variables, corresponding to a multiplicative relationship in the original variables, well-known from Cobb-Douglas type functions.

Linear specifications correspond to first order approximations and the natural next step towards a workable form is to use quadratic approximations, possibly in the log of the variables. A second order approximation using log variables gives the so-called translog form. In a cost function specification with p outputs and no prices, it pictures the relationship as

$$ln\, C_i = b_o + \sum_{j=1}^{p} b_j \, ln\, y_{ij} + \sum_{j=1}^{p}\sum_{k=1}^{p} b_{jk} \, ln\, y_{ij} \, ln\, y_{ik} + u_i + v_i$$

where $C_i$ is the total cost of the i-th unit, $y_{ij}$ is the j-th output quantity of the i-th unit, and the b's are unknown parameters to be estimated.

This form is a so-called flexible form. This means intuitively that we do not impose any unnecessary restrictions on the functional form (and more precisely on the behaviour (elasticities) that results from the function is free to begin with). The drawback of this flexibility is on the other hand that the results may be harder to interpret, that more parameters must be estimated and that the resulting estimated model may suffer from curvature violations such that the output sets for example do not have the properties of normal output sets.

The data needed to estimate a translog function is defined in part by its number of free parameters (degrees of freedom) which is 3+p+p(p+1)/2 if we work with a truncated normal distribution of inefficiencies. The 3 variables are the unknown parameters of the inefficiency and noise distributions while p+p(p+1)/2 is the number of parameters b related to the p independent variables y (we have symmetry $b_{jk}=b_{kj}$).

***Software***

In practice calculating SFA results requires the use of a software package that is not included in the standard programs installed in most personal computers. There several software packages that have dedicated procedures for SFA and COLS models. These include LIMDEP[7] and Frontier[8]. These dedicated software provide the easiest access to SFA. In addition, it is of course possible to develop the parameter and efficiency estimates using general purpose statistical software like R, S, SAS or GAMS. Since the typical estimation is based on the maximum likelihood approach the user basically needs a good non-linear optimization routine.

## 2.4 Use of SFA as analogous model

In this chapter, we briefly comment on the use of SFA as a complementary model to DEA. The section is based on the experiences the Norwegian and Austrian regulators have on SFA. However we start with more general discussion on the role of SFA in regulation.

### 2.4.1 The role of SFA

In practice there are a number of question related to the actual use of SFA in regulations. The choices related to the use are discussed shortly before discussing the actual experiences in the next subsections.

*How to use multiple models effectively in regulation?* DEA and SFA are complementary in the sense that DEA provides a stable and powerful data driven mechanism for regulation of verifiable outputs and inputs, whereas SFA provides a safeguard to data errors and stochastic influences in the cost and output data. As hinted at already in Agrell and Bogetoft (2003d), multiple models are useful only to the extent that they have a distinct application area in the overall regulatory approach. Possible uses are e.g. to provide a secondary tool to (i) provide a well-founded base for solving the super efficiency problem (in possible future DEA yardsticks) and (ii) a filter to detect, estimate and reimburse potential outliers. The selected use should be in line with the selected regulation system.

*How to address conflicting results (order, scale)?* The average SFA scores will often be higher than the average DEA scores at equivalent returns to scale, since an extra error term absorbs part of the detected distance from the frontier. This additional error may more than compensate for the added restriction on the cost structure – especially when the underlying structure is not too complex. It follows that SFA may be an overly generous/conservative regulation instrument if the inputs and outputs actually are deterministic and verifiable. When moving from a process/input-oriented regulation, which inherently may suffer from information asymmetries and data errors, to an output-oriented regulation with higher aggregation on the input side, the stochastic element may be less important than the loss of flexibility. This suggests that it is advisable to choose one approach as the primary one based on the characteristics of the regulation system.

*Handling hyper-efficient firms.* A problem in DEA based yardstick is that typically some firms have no comparators and are automatically classified as efficient. These firms are called hyper-efficient. By supplementing the variable returns to scale (VRS) DEA models with more general return to

---

7 For more information, see http://www.limdep.com/

8 Frontier software is available free of charge. For more information, see http://www.uq.edu.au/economics/cepa/frontier.htm

scale conditions (e.g. non-decreasing returns to scale (NDRS) or constant returns to scale (CRS)) or by including weight restrictions, some of this problem can be eliminated. For the remaining units, the regulator may have to turn to individual handling of their performances. SFA may be informative to set cautious norms in such settings. The implementation of such a secondary option can be either formalized – or it can be guided by regulatory discretion, negotiation or a principle of accountability.

*Information dissemination principles?* Dynamic regulation transitions require a crystal clear communication on the objectives, instruments and incentives the regulator intends to use for the upcoming periods. Given the methodological and conceptual difficulties involved in merely explaining the differences between to two models, it is often advisable to use one of the models only as an internal, yet permanent, component of the regulatory process.

*Bias correction.* The DEA cost structures are in general biased towards higher costs – and more so the smaller the data sets. This holds also for local areas less dense in observations. One way to cope with this – and avoid unequal treatment of units – is to impose some general regularity on the structure, viz. via parametric models like SFA, via weight restrictions or via enhanced possibilities to rescale. Alternatively, one can correct for bias through boot-strapping or – as a simple approximation – by using inverse shell analyses. Another approach is of course to rely on increased international cooperation – an approach that have some potential merits in a Nordic setting where the comparability is – after all – quite high, and where collaboration has previously been established, cf. also Edvardsen and Førsund (2003). On the other hand, the difficulties of creating comparable data across countries should never be under-estimated making the other alternatives, including the usage of SFA worthwhile to consider as well.

### 2.4.2 The NVE benchmarking models

Ever since the 1991 Energy Act, NVE initiated limited benchmarking exercises using key performance ratios to monitor and motivate efficiency improvements in the incumbent cost-plus regime. The culmination of this predecessor to the DEA regulation model was probably the NVE (23/1997) benchmarking software tool that was publicly distributed. However, not before the efficiency requirement was individualized did NVE synthesize the benchmarking model.

The Data Envelopment Analysis (DEA) benchmarking model of NVE has been documented in Kittelsen (1993, 1994, 1996), Kittelsen and Torgersen (1993) and NVE (1994, 1995, 1996). The model was later adjusted by including quality indicators using the so-called KILE system. Two revised models for distribution and regional transmission were presented in NVE (2001). The distribution model includes the *a priori* estimated cost of non-delivered energy, KILE, to account for quality differences among firms. The actual cost of non-delivered energy is added to the operating cost, whereas the anticipated cost is added to the exogenous variables as an indicator of operating quality.

In the development of new regulation from 2007, NVE also undertook experiments with alternative estimations of a (new) DEA model, cf. Agrell and Bogetoft(2004). The model is a total cost model with cost drivers related to energy transport (energy delivered), customer administration (number of clients), capacity provision (line length LV and HV) and some proxy for environmental complexity (expected KILE).

In particular, two SFA models based on the translog specification but using different notions of capital costs (based on book and new values, respectively) were tests. With new values, the SFA estimation gave very reasonable results while the use of capital costs based on book value lead to an ill-specified model with wrong signs on several of the costs drivers.

The correlation with the DEA results were not too impressing suggesting that the SFA model were either too restrictive in its specification of the relation between outputs and costs or that more flexible DEA were basically picking up random variations in data as inefficiency differences.

At NVE, the SFA have been used as part of the model validation process but its role in the future regulation and in particular its role in relation to the dominating primary model, the DEA model is as of now undecided. Some important issues however were identified in Agrell and Bogetoft (2004). The inevitable principal issue is how to combine (potentially conflicting) methods in a coherent regulation approach. In particular, we address the questions of (i) utilization, (ii) conflicting results and (iii) information dissemination.

In the NVE case, where frontier analysis is well established as regulatory instrument, SFA can be seen as a secondary tool to (i) provide a well-founded base for solving the super efficiency problem (in possible future DEA yardsticks) and (ii) a filter to detect, estimate and reimburse potential outliers. Both objectives contribute to the credibility and stability of the norm value, consistent with the overall strategy towards more performance based regulation outlined in Agrell and Bogetoft (2003a).

To avoid problematic interpretation of potentially *conflicting results*, Agrell and Bogetoft(2004) argued in favor of keeping DEA as the primary estimator of operator inefficiency. The results from SFA would intervene in the internal review of DEA scores to resolve outlier problems without resorting to pure regulatory discretion or low powered regimes. In this utilization, the relative ranking of firms in SFA has no bearing on the regulation, only the absolute scores.

What comes to the *information dissemination* it was concluded the in the Norwegian context it is advisable to use the SFA only as an internal, yet permanent, component of the regulatory process. Dynamic regulation transitions require a clear communication on the objectives, instruments and incentives the regulator intends to use for the upcoming periods. Given the methodological and conceptual difficulties involved in explaining the differences between DEA and SFA, it was decided that only on should be used in information dissemination.

In the Norwegian context, Agrell and Bogetoft(2004) analyzed a series of models, both parametric SFA models and non-parametric DEA models for NVE. The latter have been analyzed using traditional techniques ignoring bias and uncertainty and they have been analyzed using state of the art bootstrapping to get insight in the bias and uncertainties involved. In the CRS cost efficiency model based on 1996-1999 data, NV and global prices, the results of the three principal approaches are illustrated in Figure 2.6 below.

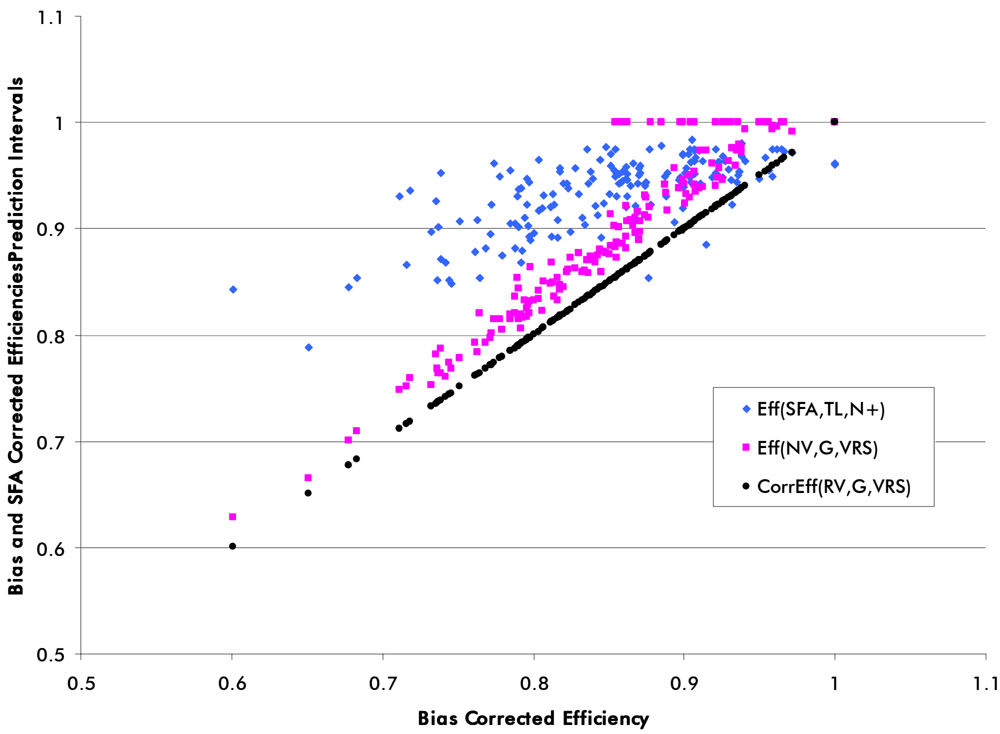



**Figure 2.6** Comparison of efficiency scores based on SFA (Eff(SFA, TL, N+)), variable returns to scale DEA (Eff(NV, G, VRS)), and bias corrected variable returns to scale DEA CorrEff(RV, G, VRS))

This comparison is interesting. It shows how the bias corrected efficiencies are (with very few exceptions) uniformly the toughest. The uncorrected are more favourable to the distribution companies, and the SFA estimates are the most favourable except for a few units, mainly those on the efficient frontier in the traditional DEA model. In a yardstick regulation, this suggests that it is fair and safe to use a SFA norm to handle the hyper-efficient units.

Also, we see that the raw DEA estimates in many cases give a reasonable compromise between the most demanding norm, the bias corrected norm, where uncertainty in the data is ignored but the bias in the DEA model is coped with, and the SFA model, where uncertainty is not ignored and where in fact large parts of the performance variation is attributed to unexplained uncertainty in the data. This gives – at least in the present application - some support to the use of the raw DEA measures in those cases.

### 2.4.3 Austria - the E-Control benchmarking models

In their attempt to introduce incentive regulation, Austrian E-Control formed in 2001, tried different models, both DEA models and COLS models. The exact role of the different models cannot be extracted from the available documents although there seem to have been a compromise where the final efficiency scores are weighted averages of the score from two DEA models and one COLS model.

The three models are closely related. The most detailed DEA model contain 5 cost drivers while the other DEA model and the COLS model contain 3 cost drivers formed by aggregating 3 of the drivers in the most detailed specification. The DEA model and the COLS model with the same specification give roughly similar average efficiency levels. This suggests that the underlying DEA model is almost linear. (A similar result was obtained in the NVE analysis in Agrell and Bogetoft (2004)).

Informal contacts with E-control suggest that they have also tested SFA models with translog functional form. However there is no public information on these results or the use of the results. The informal information also suggests that the parametric models were not used to select variables as such but rather to serve as a control method for the DEA runs much like it is planned in Finland.

The Austrian results, however, were not considered sufficiently reasonable and the results were therefore not shown to the industry nor used in the regulation.

# 3 Building the analogous model

## 3.1 Process and methods

The aim of the previous sections has been to provide a basis for the development of the analogous model. We have described the Finnish regulatory context, alternative efficiency analysis approaches, and discussed shortly the potential ways of using the alternative model in regulation.

As discussed above, it is most natural to use SFA as a basis for developing the analogous efficiency evaluation model. SFA complements DEA specifically in those areas where DEA has its weaknesses. Both SFA and DEA are also theoretically and practically well established. Hence SFA will be used as the primary method in the study. In the next chapters, the alternative efficiency model based on SFA will be discussed in detail.

The choice of a benchmarking model in a regulatory context is a multiple criteria problem. In our investigation of alternative models specifications, c.f. below, we have stressed the following four groups of criteria.

***1. Conceptual***

It is important that the model makes conceptual sense both from a theoretical and a practical point of view. The interpretation shall be easy and the properties of the model shall be natural. This contributes to the acceptance of the model in the industry and provides a safeguard against spurious models developed by data mining and without much understanding of the industry. To be more precise, this has to do with the choice of outputs that shall be natural cost drivers and with functional forms for example that have the right return to scale and curvature properties – e.g. that it is more expensive to produce more than less.

***2. Statistical***

It is of course also important to disciple the search of a good model with classical statistical tests. We seek models that have significant parameters of the right signs and that do not leave a large unexplained variation.

***3. Intuitive and experience***

Intuition and experience is a less stringent but nevertheless very important safeguard against false model specifications and the over- or under use of data to draw false conclusions. It is attractive that the models produce results that are not that different from the results one have found in other countries or related industries. Of course, in the usage of such criteria, one always the runs the risk of mistakes – we may screen away extraordinary but true results (Type 1 error) and we may go for a more common set of results based on false models (Type 2 error). The criteria shall therefore be used with caution.

One aspect of this that one will tend to be more confident in a specification of inputs and outputs that leads to comparable results in alternative estimation approaches, e.g. in the DEA and SFA model. The experiential basis of this is that when we have a bad model, SFA will see a lot of noise and therefore attribute the deviations from the frontier to noise rather than inefficiency. Efficiencies will therefore be high. DEA on the other hand does not distinguish noise and inefficiency so in a DEA estimation, the companies will look very inefficient. Therefore, too deviating results in the DEA and SFA estimations may be a sign that the model is not well-specified. However, it should be emphasized that the aim is not to generate the same results using a DEA and an SFA estimation. The aim is to find the right model. However, high correlation between the DEA and SFA results is an indication that the model specification is reasonable. It therefore also becomes an indirect success criterion.

*4. Regulatory and pragmatic*

The regulatory and pragmatic criteria perspective again call for conceptually sound, generally acceptable models as discussed above. Also, the model shall ideally be stable in the sense that it does not generate too fluctuation parameters or efficiency evaluations from one year to the next. Otherwise, the regulator will loose credibility and the companies will regard the benchmarking exercise with scepticism. Of course, one shall not choose a model simply to make the regulator's life easy, so it is important to remember that similar results is also a sign of a good model specification, cf. the intuitive criteria above. The regulatory perspective also comes into the application of the model. If the model is not good, a high powered incentive scheme for example would not be attractive since it would allocate too much risk on the firms.

The analysis process was based on the idea of estimating a number of SFA models based on different models structures, i.e. functional forms and input-output combinations. These models were analysed based on the above criteria. In the following sections, we briefly discuss the data and the series of estimations and comparisons we made to reach our recommended model. Properties of the recommended model, including the stability of the specification and the regulatory usage of it in combination with DEA, are then further discussed in section 4.

## 3.2 Model structure

The general framework used in all SFA estimations is that of a cost function

$$x_i = C(y_i) + u_i + v_i,\ i=1,\ldots,N$$

where $x_i$ is actual costs of DSO number i, $C(y_i)$ is the minimal costs of producing the output $y_i$ of DSO number i under normal circumstances. Furthermore, $u_i$ is the individual (additive) inefficiency term for DSO number i and $v_i$ is the noise (unexplained variation) in the data for DSO number i.

We have chosen to work with normal distributed noise as it is usually done in econometrics, i.e. we assume that the $v_i$ are independent, identically distributed normal $N(0, \sigma_v^2)$.

The a priori assumption about the inefficiency distribution in the population is that it is a truncated half normal distribution, i.e. the $u_i$ are independent half-normal $N_+(\mu, \sigma_u^2)$. It is important to understand that this is a very flexible starting point. The distribution of inefficiency is difficult to know a priori and the best approach is therefore to choose a flexible family. As illustrated in Figure 2.5, the truncated normal serves this purpose well.

The estimation of these models using maximum likelihood estimation is usually done in an alternative representation of the parameters, namely in terms of

$$\sigma^2 = \sigma_v^2 + \sigma_u^2$$

$$\gamma = \sigma_u^2 / \sigma^2$$

Observe that sigma squared, $\sigma^2 = \sigma_v^2 + \sigma_u^2$, is an indication of the total variation around the function C, while $\gamma$ is a measure of the relative importance of the inefficiency. The latter, $\gamma$, is between 0 and 1 with 0 corresponding to the case of no inefficiency and 1 corresponding to the case with no noise.

To determine the usual relative efficiency measures E, i.e. the measures that are comparable to the measures usually reported in the DEA study, we must – once the model have been estimated – calculate

$$C(y_i) / [C(y_i)+E(u_i \mid u_i+ v_i)]$$

i.e. we substitute the non-observable inefficiency $u_i$ by its conditional mean given the observed combined error and inefficiency term.

If the estimations are based on a log transformed numbers, like in the loglinear and the translog models, the efficiencies are determined in a slightly different way, namely as

$$1 / E(\exp(u_i) \mid u_i+v_i)$$

since in this case, the inefficiency terms $u_i$ are actually used in a multiplicative way, which after the log transformation becomes an additive term, and is assumed to follow a truncated normal distribution.

### 3.2.1 Functional forms

In the analyses we have worked with alternative specifications of the functional form. In the following we will explain these using a framework with just two outputs to simplify the formula and to avoid working with vector notation.

The simplest alternative is the *linear* specification

$$x_i = b_0 + b_1 y_{1i} + b_2 y_{2i}$$

like in a classical multiple linear regression approach.

If the deviations from the cost function (i.e. noise + inefficiency) are dependent on the size of the company it is said that there is heteroscedasticity in the data. A standard approach for dealing with this in regression analysis is to estimate a *normed linear* model

$$x_i/y_{1i} = b_0/y_{1i} + b_1 y_{1i}/y_{1i} + b_2 y_{2i}/y_{1i}$$

The purpose of this normalization is to allow the variance of the inefficiency and noise terms to increase with $y_{i1}$. There are good reasons to expect heteroscedasticity in the linear model. Specifically, the extra cost caused by inefficient management will most likely increase in absolute terms as the company becomes larger. Excessive costs of 1 mio Euro in a small company may correspond to excessive costs of 5 mio Euro in a larger company, for example. By estimating the normed model, we basically adopt the idea of having similar relative cost overruns. That is we assume that a 20% cost overrun is as likely in a small company as in a large company, not that a 1 mio Euro cost overrun is equally likely.

Another way to cope with the problem of heteroscedasticity, ie. to work with similar relative inefficiencies (and noise) as opposed to similar absolute inefficiencies (and noise), is to work with a *log-linear* (Cobb-Douglas like) specification

$$\ln x_i = b_0 + b_1 \ln y_{1i} + b_2 \ln y_{2i}$$

Lastly, as a slightly more detailed approximation, we have worked with *translog* forms

$$\ln x_i = b_0 + b_1 \ln y_{1i} + b_2 \ln y_{2i} + 0.5 b_{11} (\ln y_{1i})^2 + 0.5 b_{22} (\ln y_{2i})^2 + b_{12} \ln y_{1i} \ln y_{2i}$$

as explained in Chapter 2.

### 3.2.2 Variable choice

It is not a primary aim of this sub-project to actually determine the best specification since the idea is to make a model that can supplement the DEA analysis. Still, given the available data we have experimented with a series of possible specifications. The guiding principle in this connection has been the general criteria discussed above. Particular focus has been on specifications has been on specifications that are conceptually sound.

The classification of variables and parameters for the models is illustrated in Figure 3.1 below. With input X or controllable resources we primarily mean the costs that can be controlled within the time horizon of the model. The class of outputs Y is made of exogenous indicators for the results of the regulated task, such as typically variables related to the transportation work (energy delivered etc), capacity provision (peakload, coverage in area etc) and service provision (number of connections, customers etc). The class of structural variables Z contains parameters that may have a non-controllable influence on operating or capital costs without being differentiated as a client output. In this class we could find indicators of geography (topology, obstacles), climate (temperature, humidity, salinity), soil (type, slope, zoning) and density (sprawl, imposed feed-in locations). In short, we seek to capture the relevant and in particular the controllable costs on the input side, and the relevant services provided or cost drivers on the output side.

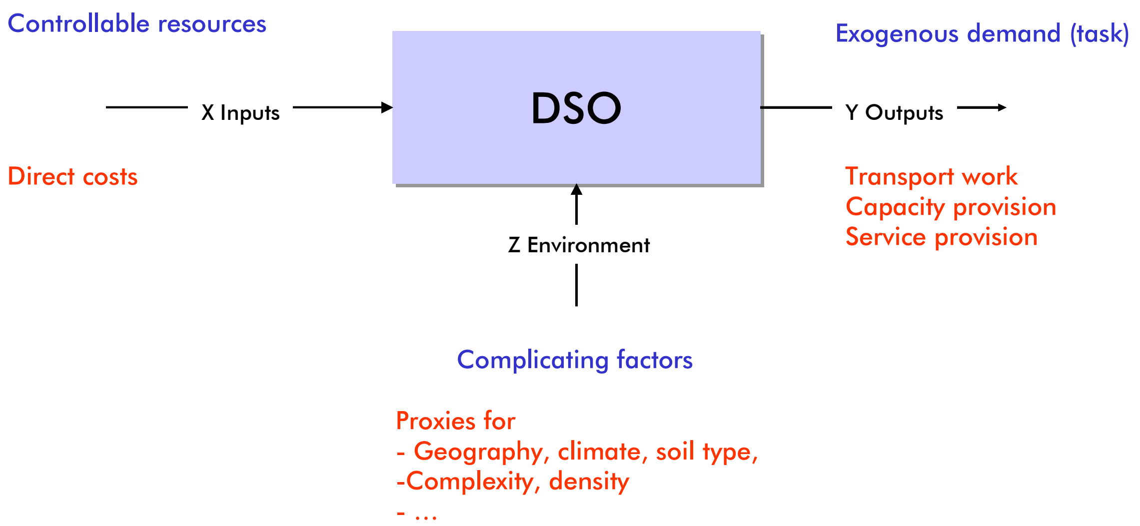



**Figure 3.1** Variable classification

The choice of variables is based on the data collected by EMV. During recent years EMV has commissioned a number of studies related to factors that can be included in the efficiency evaluation models. The starting point in this development has been the seminal study by Korhonen et al. (2000). This study presents a thorough analysis of the outputs and complicated factors. At this stage the most relevant development issues are related to the ways of taking quality and capital costs into account. In this area new data has been collected and the aim of the model development in Study A is to expand the current DEA model to better take into account quality and capital costs. The following subsections introduce the data specifications.

### 3.2.3 Input specifications

On the input side, we have experimented with different specifications. These specifications are combinations of the following three cost components.

Operational expenditure (opex) includes all the cost components that are considered to be controllable. This definition coincides with the data used in the current DEA model. The indicator includes for example personnel cost, external services, material and cost of energy used for transmission losses. Payments to transmission system operator are excluded.

Depreciation is defined on the basis of the net present value of the distribution network. As in the current regulatory system, the net present value is calculated for all the companies using the same principles; standard component prices and a range of holding times. The yearly depreciation is defined using a straight-line depreciation. At the moment this data has been classified as confidential.

Interruption costs are calculated from the indicator "average interruption time per customer" by multiplying it with the number of customers and the average cost of energy not supplied. The same average price is used for all the companies and if reflects the average customer profile of the Finnish distribution companies. In the future more detailed data on the interruptions will be available and hence the interruption costs can be calculated more precisely. For more discussion on the interruption costs and cost of energy not supplied, see the report of study A.

The starting point of the study is to take into account all the three cost components. This would reflect the aim of minimising the societal costs of electricity distribution. However to test different specifications we have tested four possible, more or less inclusive cost concepts.

X1: Opex + Depreciation + Interruption costs

X2: Opex + Depreciation

X3: Opex + Interruption costs

X4: Opex

### 3.2.4 Output specification

On the output side the study relies very much on earlier studies, especially Korhonen et al. (2000). The output set introduced in that study has been used in the current DEA model used by EMV. At the moment EMV considers that the previous studies justify the output set used and it does not see urgent development needs related to the outputs. Hence the possible extensions to the output data set are fairly limited.

In this study we have considered four possible output factors. The three first are based on the current DEA model and one additional factor is considered. The outputs are defined as follows.

Value of energy is based on the amount of energy delivered to consumption (MWh) on three different voltage levels. For each voltage level the amount of energy is multiplied by the national average distribution price (€/MWh) on that voltage level. In this way different voltage levels can be aggregated by taking into account the added value of transforming the electricity to lower voltage levels.

Number of customers refers to the number of connection points or users in the network. Numbers of customers on each voltage level are summed.

Total length of networks reflects the geographical distance of the customers from the source of energy. This is calculated by summing up the length of network on different voltage levels.

On the output side, we have experimented with different combinations of of the following variables

Y1: Value of energy

Y2: Total network length

Y3: No of customers

Y4: Replacement value

In the future it is possible that the output side could be further expanded. One possible direction could e.g. be to split the length of network into cables and overhead lines.

### 3.2.5 Environmental proxies

The basic principle in this study is that the output factors included in the model (presented above) describe the operational environment of the companies. This is justified by the analyses presented by Korhonen et al. (2000). However, it is interesting to do a comparative analysis and to see if this is valid also in the SFA context.

The impact of the environment has been dealt with via second stage analysis in this study. This means that we have estimated the models possibly without a full account for the conditions in terms of density etc under which the DSOs work. Next we have examined if the estimated variation in efficiency can be explained by some environmental proxies.

It should be noted that even though the estimation does not start with all possible environmental proxies, the output structure, i.e. the Y variables included, will typically pick up at least some of the impact of the environment. When both network length and customers are part of the outputs, for example, we also have an indirect measure of the density of the network.

Still, to control further for the impact of the environment, we have in the second stage analysis examined the explanatory power of variables like:

- Percentage of underground cables (to proxy for separate urban, semi-urban and rural companies).
- Interruption time (to proxy for difficulty)

The indicators used as environmental proxies were defined in the following way. Percentage of underground cables is calculated at the mid voltage level 6-70 kV (usually 20 kV). The length of underground cables is divided by the total length of network. This gives a proxy that can be used for separating urban, semi-urban and rural areas. Furthermore the total length of interruptions defined as the duration of interruptions at the substation level multiplied by the average number of customers per substation.

The results of the second stage analysis are presented in section 4.2.

## 3.3 Description of the data

In the development of the parametric model we have strived to make sure that all DSOs could be included in the analysis. The quality of the data seemed sufficiently high to accomplish this although we do realize that some of our evaluations, including some of the correlations and even ranking of models could be altered by more dramatic elimination of potential outliers.

In the analysis 2004 data is used as the primary data set. Unless otherwise noted we refer to this dataset. In the analysis of the stability of the results year 2003 data is used as a comparison point. However this dataset is more limited and on the input side this data includes only operational costs. Hence the comprehensive analysis of model structures etc. is based on 2004 data only.

In this section we present graphical data analysis that will support the statistical analysis used for selecting the SFA model. We first analysed the connection between total input (X1) and the output factors. This reveals the size differences of the companies play a very essential role in the dataset. Furthermore the data seem to include clear heteroscedasticity, i.e. the deviations from the assumed cost function (which correspond to inefficiency + noise in SFA) are dependent on the scale of output. This is illustrated in Figure 3.2, where heteroscedasticity causes the dataset to be showed in a conical form starting from the origin.

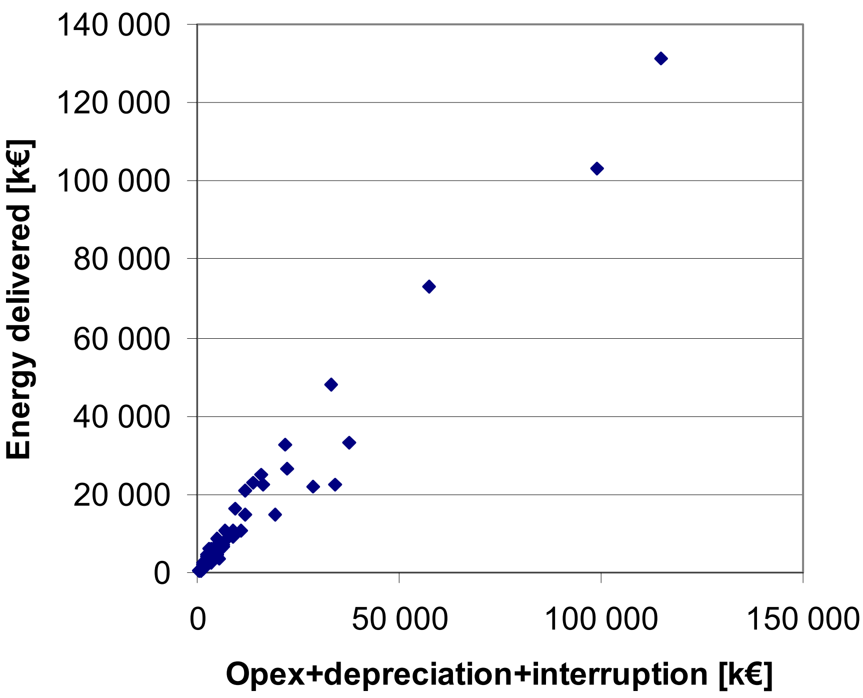



**Figure 3.2** Input versus Value of energy (Energy delivered)

Large size differences and heteroscedasticity suggest that it is better to analyse the data graphically on logarithmic scale. When both axes are transformed to logarithmic scale, straight lines that go trough the origin in the original scales are presented as straight lines in the new coordinates. This implies that data with increasing deviations from the assumed cost function (i.e. conical shape) would seem to be bound by two parallel lines and the deviations would seem to be equally large on the figure. Figures 3.3 - 3.5 present the input X1 compared to the three output indicators. These figures suggest a close to linear dependence between the input and the output factors.

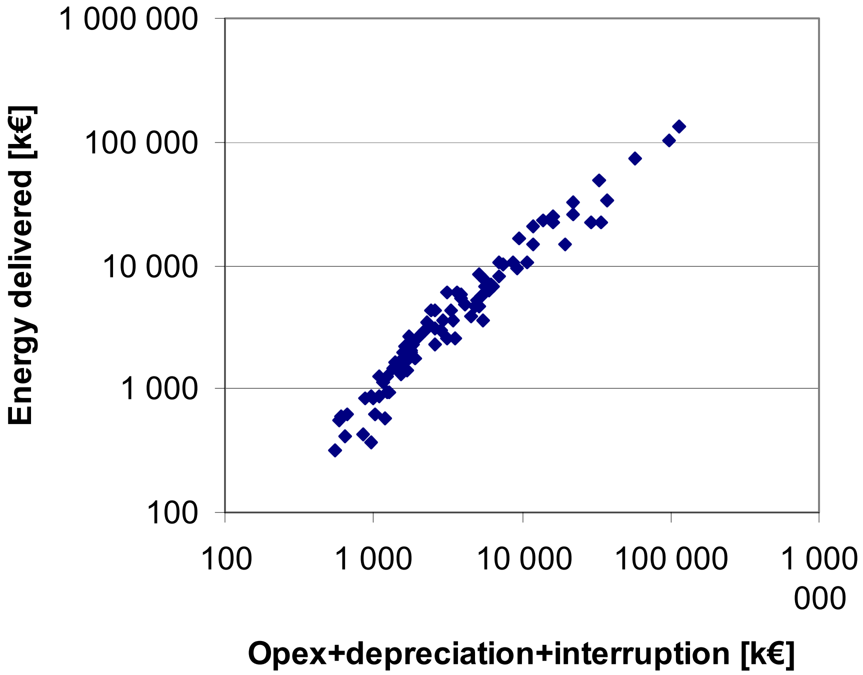



**Figure 3.3** Input versus Values of energy (Energy delivered) on logarithmic scale

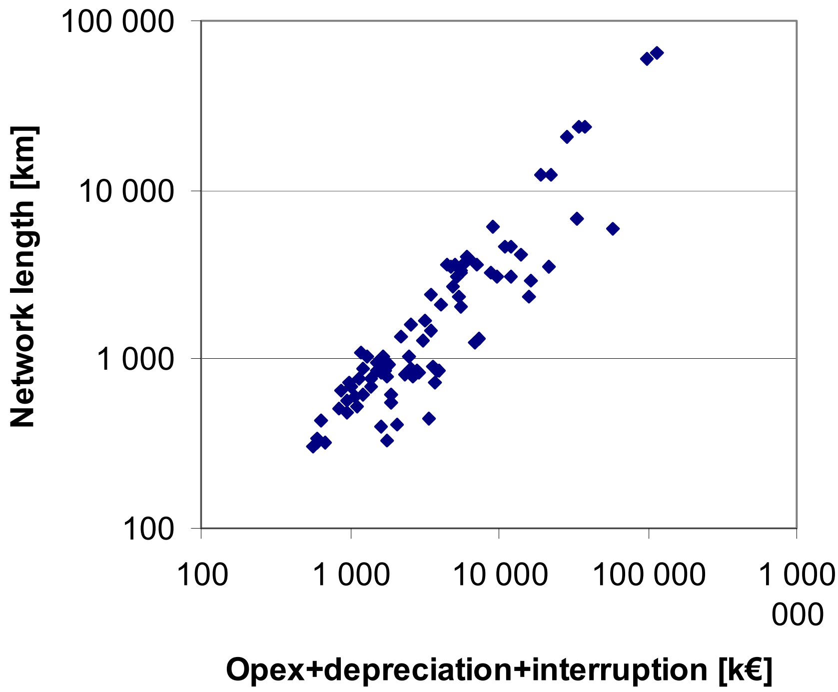



**Figure 3.4** Input versus Network length on logarithmic scale

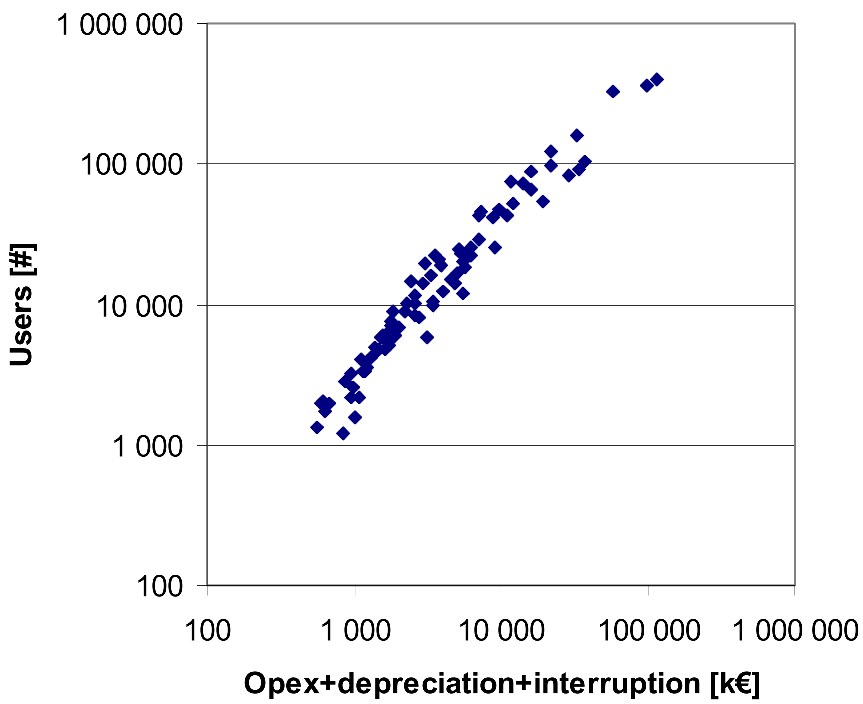



**Figure 3.5** Input versus Number of users on logarithmic scale

As heteroscedasticity is often taken into account by assuming that the error term is dependent of one of the independent factors (in this case outputs) we analysed also the dependence of input and output factors in the case where output and input factors were divided by one of the outputs. Figure 3.6 illustrates the impact of this transformation. The figure suggests that heteroscedasticity is removed by this transformation.

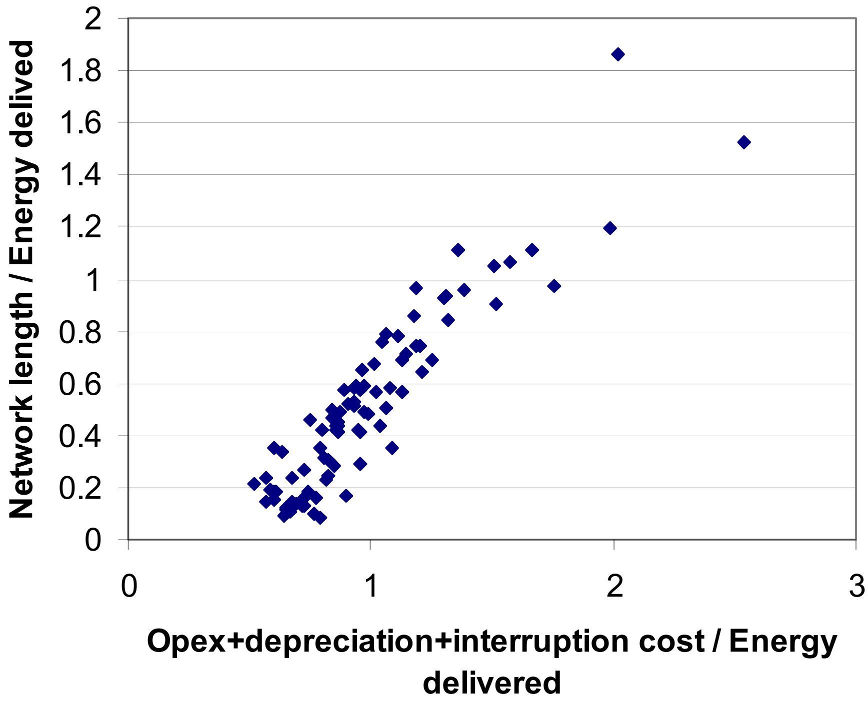



**Figure 3.6** Input v.s. network length when scaled by the value of energy

We have now looked at the link between input and the outputs. In addition to this, it is motivated to take a look at the input components. This shows that the cost components are well linked to each other and there are no outliers in the dataset. As an example, Figure 3.7 presents operational expenditure in relation to the full dimensional input (X1).

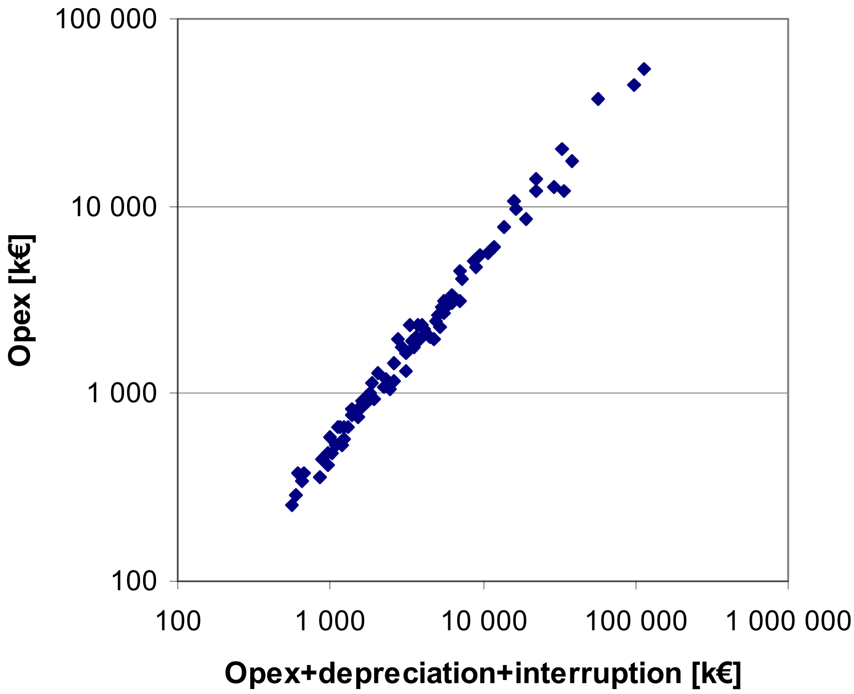



**Figure 3.7** Full input versus Opex

The above figures show that the dataset does not include any clear outliers and that there are clear dependencies between the inputs and the outputs. This provides a good starting point for the statistical analysis.

## 3.4 Selection of the model structure

We have tested a series of models structures to justify the selection of input, output and environmental factors and the selection of production technology. In practice these have been analysed simultaneously and iteratively. However, here the results are presented so that we discuss the questions related to input-output combination and functional form separately. This should make the report easier to read.

Concerning the selection of input-output combination the main purpose has been to justify the use of input-output combination independently from DEA. However, it has to be kept in mind that the dataset that is used as a starting point reflects the choices made based on the DEA analyses. This reflects the fact that full scale analysis of all the potential input and output factors has not been the goal of this study.

To justify the selection of input-output combinations 16 models structures have been investigated. These combine the above-mentioned 4 input possibilities (X1, X2, X3 and X4) with 4 output possibilities (Y1, Y2, Y3), (Y2, Y3), (Y1, Y3), and (Y1, Y4).

To make a justified selection of the functional form, a series of parametric SFA models have been developed and analysed. These functional forms have been separately estimated for each of the 16 input-output combinations. The analysed functional forms (that have been explained above) are the following.

- Linear
- Loglinear
- Translog
- Normed linear (linear with heteroscedasticity)

For the normed linear model we have analysed the use of three different output factors for norming. In the base runs, Y1 (value of energy) was used as the norming variable. After having determined the most promising specifications, a series of eight supplementary runs we experimented. These tested the use of alternative outputs, Y2 and Y3, as the norming variables.

We have for each model structure compared with a series of non-parametric DEA models. In this report we compare the results to the six most important ones. These DEA models are the following. The abbreviations used in the later text are presented in brackets.

- Variable returns to scale (d_dea_far_vrs)
- Decreasing returns to scale (d_dea_far_drs)
- Non-decreasing returns to scale (d_dea_far_ndrs)
- Constant returns to scale (d_dea_far_crs)
- Bias corrected non-decreasing returns to scale (d_dea_far_ndrs_biascorr)
- Bias corrected constant returns to scale (d_dea_far_crs_biascorr)

In the abbreviations "d" indicates that we are calculating distances (efficiency scores) and "far" is an indication that the measures are Farrell (as opposed to Shepard) measures. The first four models are the normal DEA models that only vary by the assumed return to scale. VRS is variable return to scale, DRS is decreasing return to scale (possibly disadvantages of being large), NDRS is non-decreasing return to scale (possible disadvantages of being small) and CRS is the original constant return to scale model. The bias corrected NDRS and CRS models, ndrs_biascorr and crs_biascorr, are the same models as the VRS and CRS models above, except that we now correct for the bias in the estimates using bootstrapping. In the validation phase of the project, cf. Section 4.3, we have also compared the confidence intervals of the bias corrected efficiencies with the SFA models.

The results of the 24 (16+8) test runs have been analysed both from the statistical and practical point of view. The full details are too extensive to be reported here. Detailed results have been made available to EMV in Excel files that accompany the final report.

### 3.4.1 Conclusions on input-output combinations

The analysis of possible input-output combinations was based the analysis of the 24 alternative variable combination. All the functional forms were taken into account in the analysis and in practice the analysis work related to input-output combinations and functional forms proceeded simultaneously. We first discuss the input-combinations. Based on the careful examination of the results, the following conclusions can be presented.

On the input side the full dimensional cost X1 (Opex + Depreciation + Interruption cost) works well. The use of alternative definitions would not add any value from a practical or statistical point of view. Hence we suggest the use of full dimensional cost X1 as an input. According to the preliminary results from Study A this input seems attractive in the DEA context as well.

On the output side the specification (Y1, Y2, Y3) used in the current DEA model is the most attractive. Dropping any variable would reduce the explanatory power of the models. On the other hand the use of replacement value (Y4) instead of network length does not improve the results. Hence the results support the analysis behind the current DEA model.

### 3.4.2 Summary of results concerning the functional form

The following paragraphs, figures and tables aim at summarising the key results for the recommended input-output combination (X1 (opex + depreciation + interruption cost); Y1 (Value of energy), Y2 (Network length), Y3 (No. of customers)). This section presents the results and the conclusions are presented in the next subsection.

We first present the summary of the average efficiencies in Table 3.1. This shows that the efficiency levels are mostly between 0.8 and 0.9. Only SFA linear and SFA translog deviate from this.

**Table 3.1** Summary of the average efficiencies of compared models

| MODEL | Average Efficiency |
|---|---|
| SFA Linear | 0.72 |
| SFA Loglinear | 0.88 |
| SFA Translog | 0.98 |
| SFA Normed linear | 0.89 |
| DEA VRS | 0.89 |
| DEA CRS | 0.83 |
| DEA DRS | 0.85 |
| DEA NDRS | 0.86 |
| DEA CRS Biascorrected | 0.80 |
| DEA NDRS Biascorrected | 0.82 |

We also present the estimation results for the four alternative functional forms. These are summarised in Tables 3.2 - 3.5. In these tables we give both the parameter values of the estimated functional form and the details of the estimated noise and inefficiency distributions. Recall that sigma-squared is a measure of the total variation around the best function form in the given class (linear, log linear etc), Gamma is a measure of the fraction of variation which is due to inefficiency and Mu given the mean value of the normal distribution that underlies the truncated normal distributed inefficiency distribution.

The results show there are clear problems related to the significance of the parameters in the translog model. All the co-efficients are statistically insignificant. The model also interprets almost all the variation as noise as can be seen from the low value of Gamma. The variation (sigma-squared) is the lowest of all the models. For the other specifications the individual results do not reveal anything alarming. Comparison of the results shows that in the normed linear model the share of inefficiency in the total error term is the highest. Log linear model has the second lowest error term after the translog model. The direct comparison of the total variation sigma-squared is impossible as the data has been transformed either by taking the logarithm or by dividing the data by energy value.

**Table 3.2** Estimation results for the parameters of linear model specification

| | Coef | std.err | t-ratio |
|---|---|---|---|
| Intercept | -535.8 | 180.8 | -2.964 |
| Y1 | 0.1513 | 0.0458 | 3.306 |
| Y2 | 0.6899 | 0.0241 | 28.65 |
| Y3 | 0.1212 | 0.0119 | 10.19 |
| sigma-squared | 1432577 | 1.857 | 771370 |
| Gamma | 0.6942 | 0.1035 | 6.707 |
| Mu | -226.2 | 532.8 | -0.4246 |

**Table 3.3** Estimation results for the parameters of loglinear model specification

| | Coef | std.err | t-ratio |
|---|---|---|---|
| Intercept | -0.1393 | 0.1542 | -0.9040 |
| logY1 | 0.2320 | 0.0856 | 2.710 |
| logY2 | 0.3304 | 0.0273 | 12.08 |
| logY3 | 0.4068 | 0.0857 | 4.744 |
| sigma-squared | 0.0570 | 0.0624 | 0.9132 |
| Gamma | 0.8451 | 0.1752 | 4.8246 |
| Mu | -0.1461 | 0.5516 | -0.2650 |

**Table 3.4** Estimation results for the parameters of translog model specification[9]

| | Coef | std.err | t-ratio |
|---|---|---|---|
| Intercept | 2.858 | 1 | 2.858 |
| logY1 | 1.353 | 1 | 1.353 |
| logY2 | 0.5522 | 1 | 0.5522 |
| logY3 | -1.325 | 1 | -1.325 |
| ½(logY1)^2 | 0.7069 | 1 | 0.7069 |
| logY1*logY2 | -0.1661 | 1 | -0.1661 |
| logY1*logY3 | -0.5885 | 1 | -0.5885 |
| ½(logY2)^2 | 0.3894 | 1 | 0.3894 |
| logY2*logY3 | -0.1789 | 1 | -0.1789 |
| ½(logY3)^2 | 0.8126 | 1 | 0.8126 |
| sigma-squared | 0.0108 | 1 | 0.0108 |
| Gamma | 0.05 | 1 | 0.05 |
| Mu | -1.96E-10 | 1.00E+00 | -1.96E-10 |

[9] The translog model estimation is not statistically satisfactory as can be seen by the reported standard deviations that are the same for all parameters. The degrees of freedom in the translog are quite large, in fact too large. The optimization routine does not converge well and the log likelihood obtained from a pure OLS specification (with noise but no inefficiency) is actually larger that the full SFA log likelihood for the reported parameter values.

**Table 3.5** Estimation results for the parameters of normed linear model specification

| | Coef | std.err | t-ratio |
|---|---|---|---|
| Intercept | 131.96 | 2.402 | 54.94 |
| Y1 | 0.2558 | 0.0551 | 4.642 |
| Y2 | 0.6567 | 0.0390 | 16.85 |
| Y3 | 0.0627 | 0.0166 | 3.776 |
| sigma-squared | 0.0893 | 0.0708 | 1.261 |
| Gamma | 0.9606 | 0.0395 | 24.32 |
| Mu | -0.5858 | 0.6762 | -0.8664 |

The correlations of the efficiency scores of the compared models were also analyzed. These key results are summarized in Table 3.6. We see in particular that the DEA and the translog or the simpler normed linear SFA specification have quite good correlations. As discussed earlier, this is not a criterion in is own, but an indication that the variable specification is sensible.

To get a better understanding of the actual meaning of the correlations, we can look at the differences between the efficiency scores. For example the correlation 0.80 of DEA NDRS and SFA normed linear model mean that the difference (SFA minus DEA NDRS score) varies between -0.113 and +0.118. Half of the population have difference between -0.009 and +0.070. Furthermore, the average difference is 0.035, which indicates the SFA scores are on average higher that the DEA scores. The individual differences are illustrated in Section 4.6.

**Table 3.6** Correlation of the efficiency scores

| | d_dea_far_vrs | d_dea_far_drs | d_dea_far_ndrs | d_dea_far_crs | d_sfa_linear_far | d_sfa_loglinear_far | d_sfa_translog_far | d_sfa_normedlinear_far |
|---|---|---|---|---|---|---|---|---|
| d_dea_far_vrs | 1 | 0.8353 | 0.8401 | 0.7546 | 0.2004 | 0.3656 | 0.7080 | 0.6337 |
| d_dea_far_drs | 0.8353 | 1 | 0.5944 | 0.8842 | 0.5499 | 0.4131 | 0.6551 | 0.4971 |
| d_dea_far_ndrs | 0.8401 | 0.5944 | 1 | 0.7798 | -0.0607 | 0.4904 | 0.8429 | 0.8010 |
| d_dea_far_crs | 0.7546 | 0.8842 | 0.7798 | 1 | 0.3865 | 0.5704 | 0.8371 | 0.6881 |
| d_sfa_linear_far | 0.2004 | 0.5499 | -0.0607 | 0.3865 | 1 | 0.2185 | 0.0727 | -0.0262 |
| d_sfa_loglinear_far | 0.3656 | 0.4131 | 0.4904 | 0.5704 | 0.2185 | 1 | 0.6557 | 0.8546 |
| d_sfa_translog_far | 0.7080 | 0.6551 | 0.8429 | 0.8371 | 0.0727 | 0.6557 | 1 | 0.8388 |
| d_sfa_normedlinear_far | 0.6337 | 0.4971 | 0.8010 | 0.6881 | -0.0262 | 0.8546 | 0.8388 | 1 |

Furthermore, the results suggest that the efficiency distribution should be related to the size. This means that suitable parametric models are loglinear, translog model and normed linear model. The analysis concerning various norming variables in the normed linear model suggest the use of value of energy (Y1) as the indicator of company size. Energy delivered is also a cost driver (or related to a cost driver) by itself

### 3.4.3 Conclusions on the functional form

On the basis of the analysis of different input-output combinations and functional forms, we can present the following conclusions concerning the analysed functional forms. Each functional form is first analysed based on the criteria for selecting the model that are presented above and then we present out conclusion based on these results.

***Linear***

Conceptually linear functional form is attractive. It is very easy to interpret. However it is possibly the functional form is too simple to describe the real cost structure. Statistically the model clearly suffers from heteroscedasticity as the absolute inefficiency (in €) depends on the size of the company. This also leads to a situation where large companies get higher efficiency score that small ones and this is contradictory with the experience that we have e.g. constant returns to scale DEA. Furthermore the results have very low correlation with DEA results.

The problems with the heteroscedasticity is the most serious problem with this model and we cannot recommend using it.

***Loglinear***

Compared to linear model loglinear model is slightly more complicated as the data is transformed by taking a logarithm of both the inputs and outputs. However it is still relatively easy to understand and interpret. The transformation solves the problem of heteroscedasticity as illustrated by the figures based on the data. Furthermore the model gives a direct indication of the returns to scale properties. In this case the function is very close to constant return to scale (or only very slightly increasing returns to scale).

However there is a significant conceptual problem with the model. As we are estimating a cost function, the production possibility set covered by the function is not convex, In other words, the iso-cost curves (which are straight lines in the log-log space) are such that a linear combination of two points on this curve is outside the production possibility set. Possibly due to this conceptual difference, the results have fairly low correlation with DEA efficiency scores.

The conceptual problems related to the loglinear cost function are illustrated in Figure 3.8. The picture on the left presents an imaginary data set (red points) and a straight line that corresponds to an iso-cost curve that is produced by fitting a linear model in this type of data. The right hand side presents the same iso-cost curve in the original scale and shows how a DEA frontier fitted in the imaginary data set would look like (red line).

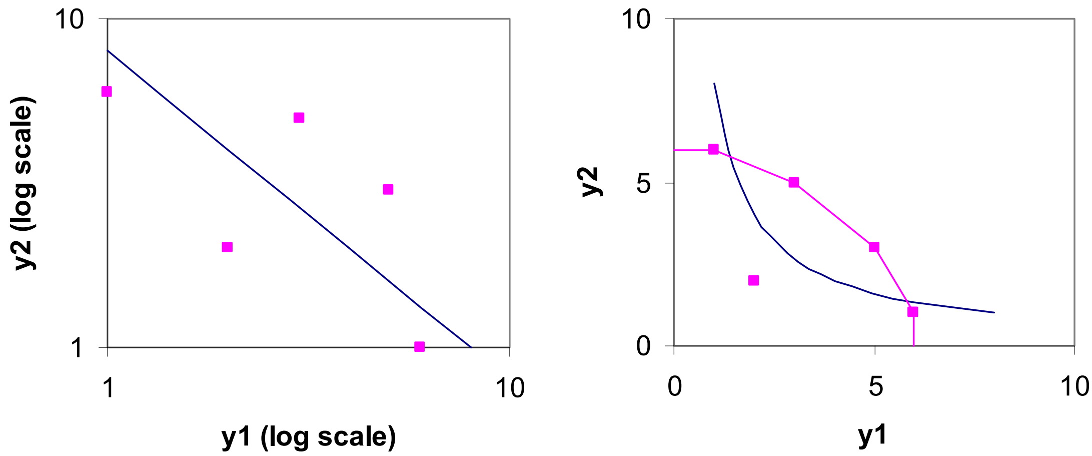



**Figure 3.8** Illustration of loglinear cost function

Due to the conceptual problems related to the loglinear model we suggest discarding this model.

### *Translog*

Translog model is a flexible functional form and is often used as a starting point in building a model. Conceptually is it attractive as it gives a second order approximation of the production function in stead of just linear. However, the flexibility leads to fairly large number of parameter and the interpretation of the model is quite difficult. As log linear model this approach may lead to non-convex production possibility set.

In this case the translog model shows a very good fit to the dataset, but almost all the coefficients are statistically insignificant. This suggests that there is too much flexibility in the model. Also form an intuitive and experience point of view average efficiency of 0.98 with the suggested inputs and outputs seem too high. On the other hand we see high correlation with the DEA results.

The results suggest that translog is too flexible, it has clear problems in the interpretation and most of the parameters are insignificant. Hence, it is not suitable.

### *Normed linear*

Conceptually the normed linear model is similar to the linear model – it is easy to interpret and has natural properties. It solves the problems related to heteroscedasticity without the conceptual problems related to the loglinear model. On the other hand, it has the same potential problem of being too simplistic as the linear model. However statistically the model gives good results, and all the coefficients are significant. The results also seem to be in line with the expected level of efficiency – the efficiency scores are slightly higher that in DEA. There is also high correlation with DEA results, especially NDRS DEA model.

Based on this normed linear model seems to be the most suitable functional form.

## 3.5 Recommended model structure

Based on our analysis on the input-output combinations and functional forms, our initial recommendation is to use normed linear model with opex+depreciation+interruption cost (X1) as inputs and value of energy (Y1), network length (Y2), and number of customers (Y3) as outputs. In the suggested model suggested normalisation factor is the value of energy (Y1).

The estimated cost function can be summarised as follows[10]:

Total cost = 132 + 0.26 Energy value + 0.66 Network km + 0.06 Customers

The total cost in this model is presented in thousand euro and the interpretation is that there is an initial cost of 132 thousand euro for all the companies and after that each € of energy value increases the annual costs with 26 cents, each kilometre of line with 660 euro and each customer with 60 euro.

The analysis of the noise and inefficiency results shows that the inefficiency scores are based on a truncated normal distribution with negative mean[11] and hence efficiency scores close to 1 are more frequent than lower efficiency scores. Most of the total variation is classified as inefficiency. According to the estimation results, the share of inefficiency in the total error term is actually very high, which suggests a good model structure and quality of data.

### 3.5.1 Returns to scale assumption of the recommended model

As the purpose of this study has been to produce result that are comparable to the DEA models analysed in study A, it is important to analyse the returns to scale properties of the recommended model in more detail. The returns to scale assumption of the DEA model is of specific interested and this question was analysed separately also in SFA.

In the recommended SFA model presented above, no a priori assumptions on the returns to scale properties were made. The recommended model includes a positive constant term, i.e. a positive initial cost. This means that the model corresponds to increasing returns to scale assumption. In the DEA terminology this is called non-decreasing returns to scale (NDRS). In the following the acronym NDRS is used for the above presented SFA model that includes the constant term.

The parametric structure of SFA allows some control over the returns to scale properties. In the case linear functional form, the constant term defines the returns to scale properties. Hence setting it to zero in the estimation will lead to a model that corresponds to the constant returns to scale (CRS) DEA model. In the following we compare this CRS SFA model to the above presented NDRS model.

The estimation results related to normed linear CRS model are presented in Table 3.7.

[10] Detailed results were presented in Table 3.5

[11] The exact mean is -0.586*Energy, see illustration of distributions with negative mu in Figure 2.5

**Table 3.7** Estimation results for the parameters of normed linear CRS model specification

| | Coef | std.err | t-ratio |
|---|---|---|---|
| Y1 | 0.2673 | 0.0522 | 5.124 |
| Y2 | 0.7845 | 0.0309 | 25.40 |
| Y3 | 0.0565 | 0.0169 | 3.337 |
| sigma-squared | 0.1192 | 0.0650 | 1.835 |
| Gamma | 0.9693 | 0.0238 | 40.66 |
| Mu | -0.6798 | 0.4563 | -1.490 |

The constant term in the above presented normed linear NDRS model is statistically significant and hence it cannot be dropped from the model with statistical reasons. On the other hand the results of the loglinear model suggest close to constant returns. Also from a regulatory point of view, constant returns to scale model has some desirable properties. Most significantly it is neutral in terms of the scale of operation and hence the returns to scale assumption does not *per se* direct the structural development of the industry. This means that companies are directed towards the most economic scale (for example through mergers, acquisitions etc.)

When we estimate the normed linear CRS model we get the following results:

Total cost = 0.27 Energy value + 0.78 Network km + 0.06 Customers

Compared to the normed linear NDRS especially the impact of length of network has increased is now 780 €/km compared to the earlier 660 €/km. All the coefficients are significant also in this model.

We have also analysed the correlation of the SFA CRS model the most important DEA models. These results are presented in Table 3.8.

**Table 3.8** Correlation of SFA normed linear CRS and NDRS models with DEA models

| | SFA normed linear | |
|---|---|---|
| | CRS | NDRS |
| d_dea_far_vrs | 0.59 | 0.63 |
| d_dea_far_drs | 0.74 | 0.50 |
| d_dea_far_ndrs | 0.55 | 0.80 |
| d_dea_far_crs | 0.77 | 0.69 |

When we compare the normed linear CRS model to the DEA models, we se, that the correlation with DEA CRS results increase from 0.69 to 0.77. On the other hand the correlation with DEA NDRS drops from 0.80 to 0.55.

Compared to the normed linear NDRS model the average efficiency drops slightly, from 0.89 to 0.88.[12] On the unit level changes are of course larger. Figure 3.9 illustrates the difference in the efficiency score relative to the input. This shows that the use of Normed linear NDRS model is more favourable to the small companies.

[12] The average efficiency scores of the main models compared are preseted in Table 3.1

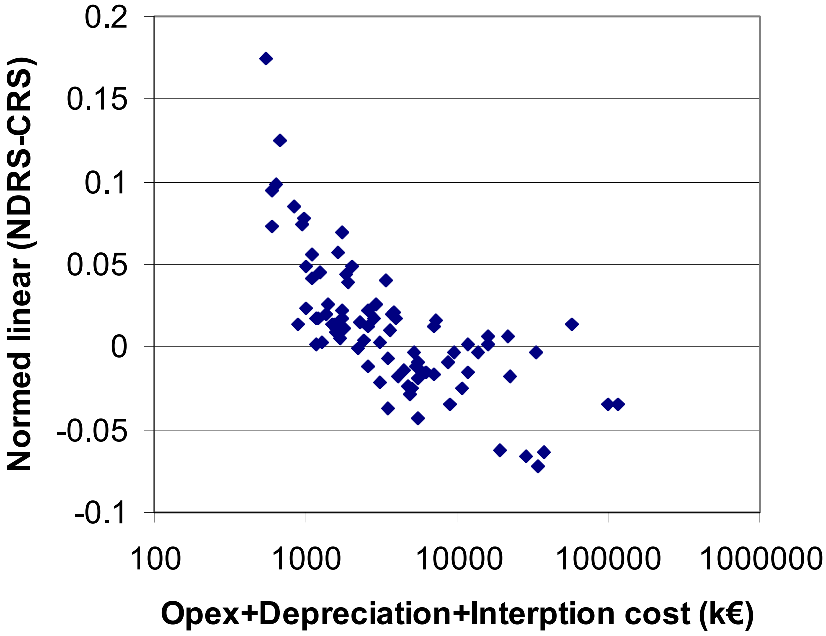



**Figure 3.9** The difference between Normed linear NDRS and CRS compared to the size of the company

We can conclude that both the normed linear NDRS and CRS model are applicable and well motivated. The choice depends on the intended use and assumptions that are found desirable when deciding on the regulatory system.

### 3.5.2 Conclusion

Based on our analysis we recommend the use of normed linear SFA model. The models can be estimated either in CRS or NDRS form. The analysis does not give any clear answer for the choice between these two. Results suggest either constant or slightly increasing returns to scale. Statistical results suggest that the NDRS version would be the right choice. On the other hand, CRS model is more neutral from the regulatory point of view. When the results are used together with DEA results, it is important to use the corresponding returns to scale assumptions in both methods.

# 4 Analysis of regulatory implications

## 4.1 Stability over years

From the regulatory point of view it is important to analyse how stable the results are form year to year. It would have been best to analyse the changes using the model specification recommended above but unfortunately this was not possible. Year 2003 data does not include the information on depreciation, and on the other hand data from quite many companies are missing from the preliminary 2005 data. Hence it was decided that the stability will be analysed based on 2003 and 2004 data sets using the inputs and outputs used in the existing DEA model. This means that instead of total cost (opex+depreciation+interruption cost) only operational expenditure (opex) is used as an input.

Table 4.1 presents the estimation results. This shows that the coefficients of the model change only slightly from year to year. This suggests that the efficiency scores would be fairly stable from year to year.

**Table 4.1** Estimation results for 2003 and 2004 data

| **2003** | **sfa.normedlinear.vrs** | | | **2004** | **sfa.normedlinear.vrs** | | |
|---|---|---|---|---|---|---|---|
| Output | coef | std.err | t-ratio | Output | coef | std.err | t-ratio |
| Constant | 88.70 | 3.806 | 23.30 | Constant | 89.60 | 1.011 | 88.62 |
| Energy | 0.1935 | 0.0380 | 5.089 | Energy | 0.1828 | 0.0411 | 4.446 |
| Network | 0.2069 | 0.0245 | 8.435 | Network | 0.2264 | 0.0248 | 9.116 |
| Users | 0.0320 | 0.0119 | 2.686 | Users | 0.0327 | 0.0131 | 2.491 |
| sigma-squared | 0.0284 | 0.0118 | 2.401 | sigma-squared | 0.0223 | 0.0067 | 3.316 |
| Gamma | 0.9076 | 0.0597 | 15.20 | gamma | 0.8704 | 0.0645 | 13.49 |
| mu | -0.3213 | 0.1898 | -1.693 | mu | -0.2788 | 0.1225 | -2.277 |

| **2003** | **sfa.normedlinear.crs** | | | **2004** | **sfa.normedlinear.crs** | | |
|---|---|---|---|---|---|---|---|
| Output | coef | std.err | t-ratio | Output | coef | std.err | t-ratio |
| Energy | 0.1731 | 0.0375 | 4.611 | Energy | 0.1575 | 0.0668 | 2.358 |
| Network | 0.2842 | 0.0290 | 9.810 | Network | 0.3100 | 0.0328 | 9.446 |
| Users | 0.0341 | 0.0113 | 3.013 | Users | 0.0357 | 0.0155 | 2.297 |
| sigma-squared | 0.0476 | 0.0321 | 1.483 | sigma-squared | 0.0205 | 0.0409 | 0.5022 |
| Gamma | 0.9460 | 0.0494 | 19.14 | gamma | 0.8118 | 0.3284 | 2.472 |
| mu | -0.4245 | 0.4087 | -1.0385 | mu | -0.1156 | 0.7237 | -0.1560 |

When we analyse the change in efficiency scores, we see that for many companies the efficiency scores have changed more than 5 % points. Hence it was analysed what causes the changes. It turned out that changes in opex are the most significant source of changes. Figures 5.1 and 5.2 illustrate the connection between change in opex and change in efficiency score.

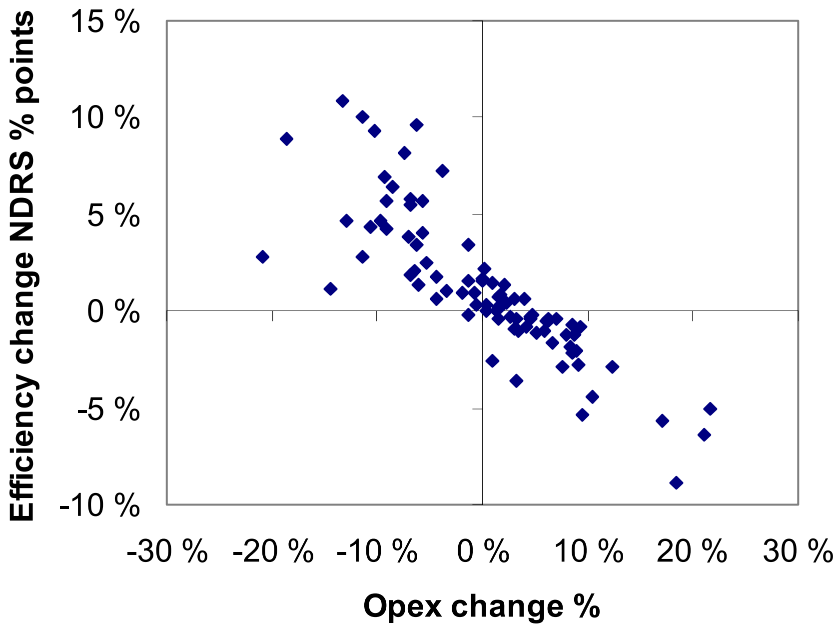



**Figure 4.1** Change in normed linear NDRS efficiency score compared to change in opex

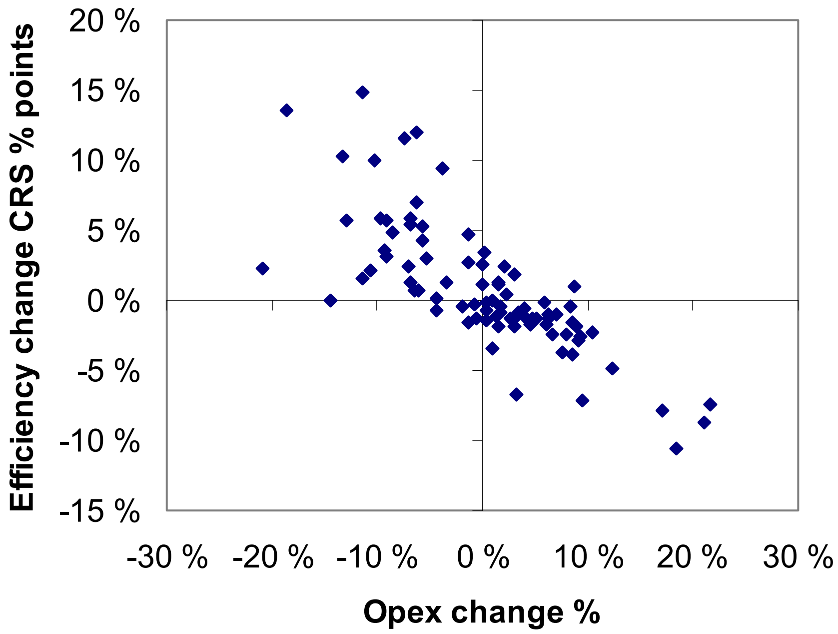



**Figure 4.2** Change in normed linear CRS efficiency score compared to change in opex

To better understand the actual change in the production function (frontier) also the shift in the frontier was calculated for each company. The changes in the frontier vary from -2.3% to +3.4% for the normed linear NDRS model and from -3.1% to +4.6 for the normed linear CRS model. Hence the NDRS model seems to be slightly more stable. However, it has to be noticed that the changes have not been analysed based on the model that was recommended and only two years were included. The frontier shifts are illustrated in figures 4.3 and 4.4.

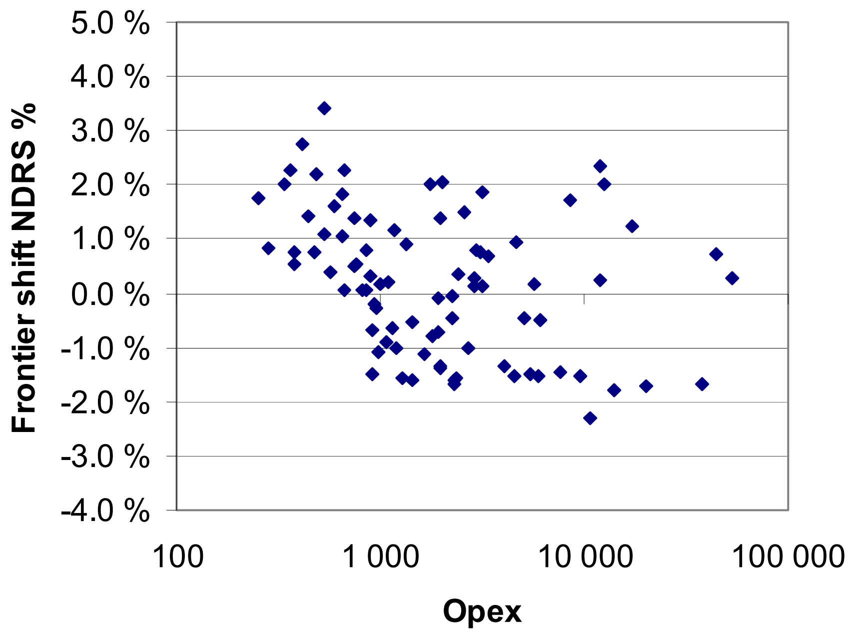



**Figure 4.3** Frontier shift in normed linear NDRS compared to opex in 2004

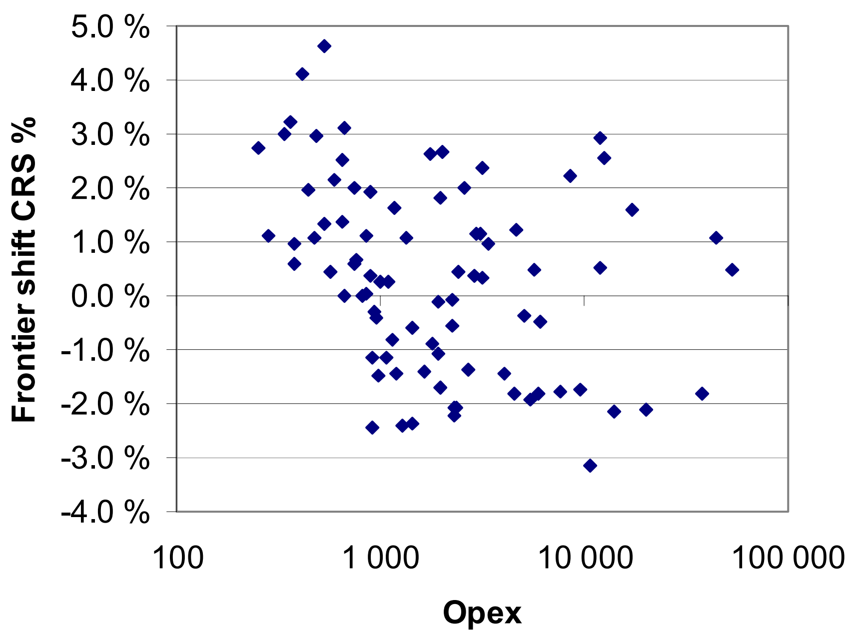



**Figure 4.4** Frontier shift in normed linear CRS compared to opex in 2004

Based on these results, we can conclude that the frontier is fairly stable and gives a stable comparison point. The frontier is more stable than in DEA. Efficiency changes are mostly caused by changes in the inputs and outputs of the companies. The results suggest that use more than one year data could be used to smooth out fluctuation when setting the improvement target for regulatory period of four years.

## 4.2 Potential biases

To analyse the potential biases of the SFA model we compared the efficiency scores to the following indicators:

- Input X1 as a proxy for the size of the company
- Percentage of cabling (6-70 kV network) as a proxy to separate between urban and rural companies
- Ratio of opex and depreciation as a proxy for different strategies concerning level of investments in the network.
- Interruption cost per customer as a proxy for external conditions causing interruptions

As it has been agreed that the efficiency score of individual companies will not be published at this stage, the results cannot be illustrated by presenting figures where these indicators are presented against the efficiency scores. The results reveal that there are no alarming biases, but some observations can be made.

- Normed linear NDRS is slightly favourable for the very smallest companies. According to a second stage regression based on input X1 and the efficiency scores this bias is statistically insignificant.
- Normed linear CRS is slightly more favourable for larger companies. This dependence is also statistically insignificant in second stage regression.
- A second stage regression shows that both model versions are slightly favourable for rural companies. The efficiency scores drop by about 1 percentage points when the level of cabling increases by 10 percentage points. Although the number of companies with high percentage of underground cabled is relatively low, this dependence is statistically significant. However, the current dataset does not provide good tools for solving this issue. In the future the possibility of splitting the length of network in to two separate outputs – overhead lines and cables could be analysed.

There is no significant difference between Normed linear NDRS and CRS versions in terms of bias – both models are equally good. However the decision on the returns to scale assumption has to be made based on the intended use of the SFA model and it has to take into account the returns to scale assumption made in DEA.

## 4.3 Confidence intervals

In this section we shortly discuss the possibilities related to confidence intervals of efficiency scores and bias corrected DEA scores. The idea behind these approaches is that statistical estimation methods allow us to analyse uncertainly that is related to the results.

If we first consider the SFA results, it is theoretically possible to calculate the confidence intervals for the efficiency scores on a wanted confidence level (e.g. 95 %). This would mean that instead of the best estimate (which has been used above), we would have an interval around it. This would reflect the uncertainty related to the scores.

In DEA bias correction works in the way that it decreases the DEA tendency to give high efficiency scores to extreme units (with few or no peer units). Hence the bias corrected efficiency scores are always lower than the normal DEA scores. Furthermore, it is also possible to calculate confidence intervals for the bias corrected DEA scores.

If we consider the use of confidence intervals in regulation, the key challenge is that these could be used in the way that it benefits either the customers or the companies. Confidence intervals are linked to a question of risk sharing between customer and companies when the result includes random error. From a statistical point of view, there is no reason to adjust the results as they represent the best guess. As there is no clear reason to favour either side, the best estimate results are recommended. The use of intervals would also make the results and the link between the results and regulation more difficult to understand.

From a purely statistical point of view it is recommended that bias correction is used in DEA. This means that companies are treated more equally. However, the impact is relatively small and we loose the easy interpretation i.e. comparison to a real target, and it is more complicated to calculate the results. As standard DEA is more conservative i.e. it gives more easily achievable targets. Hence our conclusion is that bias correction would not provide clear value added in regulation.

## 4.4 Marginal impacts

In DEA the analysis of marginal impacts is interesting as some input and output factors may have no impact on the efficiency score. As the SFA model that has been recommended is based on a linear cost function, there is a straight forward relation ship between improvements in the (total) cost and the efficiency score.

It has to be noticed that as SFA includes the noise term that is estimated for each company during the estimation of the model parameters, the model should be re-estimated every time the data changes. However we can approximate the impact by assuming the change in the noise term $v_i$ is small when we make marginal changes in the data for one unit at a time. This implies that the new efficiency score of unit i after a cost change can be calculated as follows:

$$E_{i,new} = C(y_i) \;/\; (x_{i,new} - v_i),$$

where E refers to the efficiency score, and the subscript "new" refers to the new values after the changes in the cost.

As the noise term can be calculated from the current cost level, frontier cost and the efficiency score (i.e. $v_i = x_i - C(y_i)/E_i$) the formula can be rewritten as follows.

$$E_{i,new} = C(y_i) \;/\; (x_{i,new} - x_i - C(y_i)/E_i),$$

This way a 1% change in the cost leads to 0.76 – 0.96 %-point change in the efficiency score. The variance depends on the initial efficiency of the unit ($E_i$), i.e. for inefficient units larger relative changes are needed.

The approximate change in efficiency caused by the change in the output factor $y_i$ can also be calculated with the above formula. In this case the estimated frontier cost $C(y_i)$ in the nominator of the formula is replaced by the frontier cost based on the adjusted output values $C(y_{i,new})$.

## 4.5 From efficiency score to efficiency improvement target

From the regulatory point of view it is very essential how efficiency scores are transformed to improvement targets. The project does not aim at developing a detailed model for transforming the efficiency scores into improvement targets (or X-factors). However, this section discusses aspects that has to be taken into account and discusses potential approaches on conceptual level.

The starting point is that all the cost components are subject to improvement target, but they are different in terms of their controllability.

- Opex – most easily controllable in the short run
- Depreciation – affected only through changes in the network (present value) – slow changes
- Interruption cost – can be affected both by operational activities and investments.

It can be evaluated that the time for removing inefficiency i.e. catch-up speed is at least 5 years for operational expenditure and at least 20 years for capital expenditure or depreciation. As interruptions are affected both by the structure of the networks and operational practices, the time for lowering interruption costs to target level is between these two extremes.

We recommend that the following principles are taken into account when setting the improvement targets:

- Although there may in practice be different improvement potentials related to the different cost component, it is impossible to say anything detailed about the long term improvement potentials related to depreciation and opex separately. This would require defining the level of opex and depreciation that lead to the minimum total cost. Shedding light on this would require analysis of the three cost components separately.
- As only one efficiency score is calculated based on three input components, it is natural to apply the same long term improvement target to all the components and that the actual improvement target is set for the total cost.
- Different possibilities i.e. different catch-up speeds related to the components need to be taken into account. This would ideally take into account the individual situation of each company.

It seems that the ratio of depreciation and opex is the only possible indicator that could be used for differentiating between the different catch-up possibilities of each company. The ratio is to some extent dependent on environment (e.g. level of urbanisation) but is also depend on the decision of the company. It seems that calculating company specific catch-up speed and hence company specific improvement targets based on the ratio would be the most natural approach.

One possible approach would be to calculate an individual catch-up factor for each company based on the actual opex-depreciation ratio (that varies between 0.8 – 2.5). The catch-up speed could be calculated using the following principle.

catch-up per year = (opex/(opex+depreciation)*((1-efficiency)/opex catch-up time) +
(depreciation/(opex+depreciation))*( 1-efficiency)/depreciation catch-up time).

If we assume the above catch up speeds, an efficiency score of 0.8 would lead to the following improvement targets:

- Ratio 2/1 -> goal 3%-points per year
- Ratio 1/1 -> goal 2.5%-points per year

Before setting the improvement targets the role of interruption costs in the total cost and the catch-up speed related to interruption cost should be analysed. Furthermore possible ways of setting improvement targets and their impact on the regulatory system as a whole should be analysed thoroughly.

## 4.6 Ways to combine SFA and DEA

As the aim is to use the SFA results in parallel with the new DEA model, it is necessary to discuss the ways of combining the use of the models in the regulatory model. As related to setting the improvement target this report does not aim at providing a comprehensive model. As discussed in section 2, there are many possible ways of using the SFA results.

The starting point of this study has been that the SFA scores will be used in the actual regulatory model. This requires that these can be applied in the same way as the DEA scores. If the models will be used in parallel, is necessary to analyse different ways of combining SFA and DEA scores.

Before analysing the ways of combining the results, it is good to have a look at the differences between the efficiency scores. Figures 4.5 and 4.6 illustrate the differences. The results clearly show that for most of the companies SFA gives a higher efficiency score.

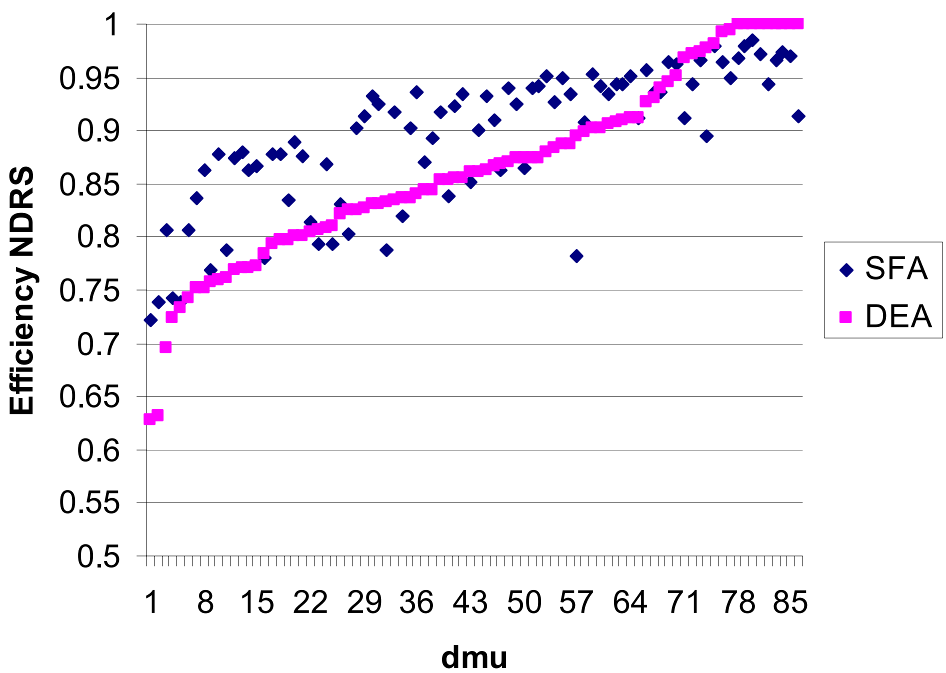



**Figure 4.5** Comparison of SFA and DEA efficiency scores, NDRS case

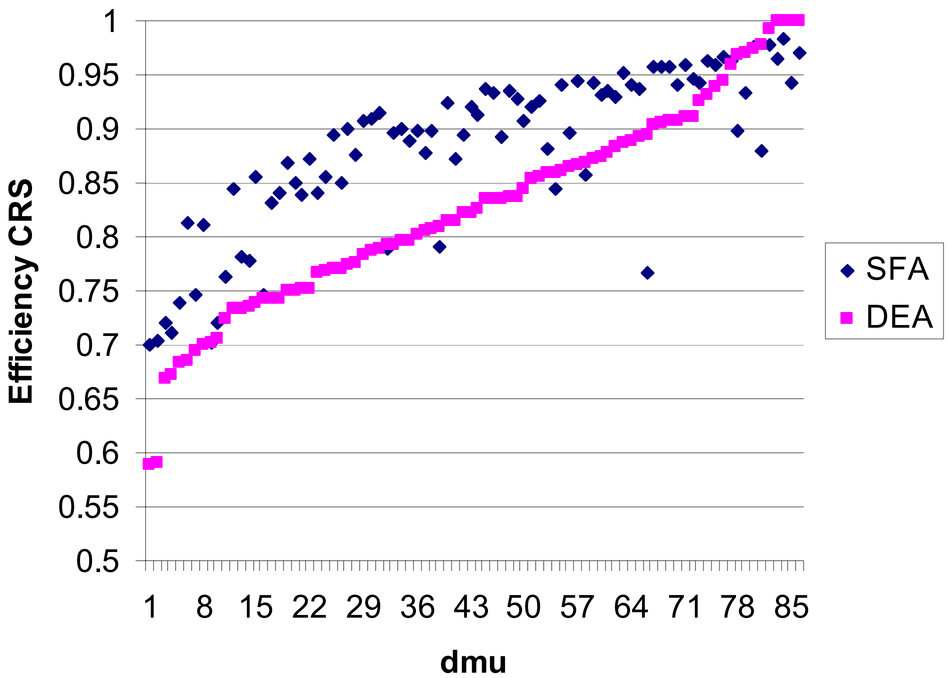



**Figure 4.6** Comparison of SFA and DEA efficiency scores, CRS case

When we consider simple ways of combining the efficiency scores, we can come up with three obvious ways: taking the higher score, taking the lower score or using the average. The following observations can be made related to these three approaches.

1. Taking the maximum of DEA and SFA scores

- If one of the score is too high (by mistake), this method leads to a situation where this biased score is used. For example, this approach does not solve the problem of hyper efficient extreme units in DEA.
- It is favourable to the companies as the smaller improvement target is chosen.
- The improvement target will be based on SFA for most of the companies.

2. Taking the minimum of DEA and SFA score

- Leads to tougher improvement targets than the used of one method only.
- If one of the score is too low by mistake, this punishes the company. The target may be unrealistic. On the other hand, the model is most favourable for the customers.
- The improvement target will be based on DEA for most of the companies.

3. Taking the average of DEA and SFA

- Filters out potential mistakes related to both approaches and filter out extremes, although does not exclude biased results totally.
- The use of average will, on average, lead to lower improvement targets than the use of DEA alone.

Also the total improvement potential of all the DSOs can be analysed. The actual impact on cost levels is of course impossible to know without knowing the actual regulatory model and the way companies respond to this. However, this analysis gives some indication of the improvement potential the models imply. The results are presented in Table 4.2. These numbers can be compared to the total input of all the companies, 764 million €. The table shows the natural result that higher average efficiency scores indicate smaller industry level improvement potential. For example for the SFA NDRS model the improvement potential is estimated to be 12% of the total input and for Max(SFA, DEA) the potential corresponds to 9%.

**Table 4.2** Total improvement target (1000 €) of all the DSOs based on different efficiency scores and their combinations

| | NDRS | CRS |
|---|---|---|
| SFA | 93 600 | 79 900 |
| DEA | 114 000 | 117 000 |
| Max(SFA, DEA) | 84 200 | 69 900 |
| Average(SFA, DEA) | 104 000 | 98 600 |
| Min(SFA, DEA) | 123 000 | 127 000 |

# 5 Conclusion and recommendations

This study has aimed at developing an alternative efficiency analysis model for analysing the efficiency improvement potential of Finnish DSOs. The purpose has been to develop a model that is analogical to the new DEA model that has been developed in a parallel study. The primary objective has been to overcome the possible estimation biases that are present in DEA. Based on a literature review, Stochastic Frontier Analysis (SFA) was chosen as a starting point. It provides a complementary approach that tackles the weaknesses of DEA.

The study analysed a number of different functional forms and input-output combinations. The models were analysed from different perspectives, taking into account conceptual, statistical, regulatory and pragmatic criteria and the earlier experience on the efficiency levels. The data set was identical to the parallel DEA study which allows direct comparison of the results. Year 2004 data has been the primary dataset. Stability of the results was analysed by comparing the results to year 2003.

The main focus of the study has been to develop the model concept, i.e. to choose the input and output factors, functional form and assumptions concerning the efficiency and noise distributions. The study clearly shows the applicability of SFA in the Finnish regulatory context. However, the exact model parameters and especially company specific efficiency scores should be re-estimated and analysed more carefully before using the model in actual regulation. The estimation should ideally be based on the actual data definitions that will be used in the regulation. For example for year 2005, new interruption data will be available.

Based on a thorough analysis of various possible models, we recommend that the efficiency analysis would be based on the use of operational expenditure, depreciation and interruption costs as an input and value of energy, network length and number of customers as outputs. Furthermore, we recommend the use of linear functional form, where heteroscedasticity is taken into account in the estimation by normalising (dividing) the input and output factors by the value of energy. Depending on the returns to scale assumptions made in DEA, the model can be estimated either with or without constant leading to constant returns to scale or non-decreasing returns to scale situation.

The analysis of regulatory consequences shows that there are no clear biases in the SFA scores. It treats different sized companies and both urban and rural companies reasonably fairly. The SFA scores are higher than the corresponding DEA scores for most of the companies. Furthermore, the stability of the SFA production function is good, providing a more stable comparison point than DEA. The results indicate that the observed changes in the efficiency scores are mostly caused by changes in the own inputs and outputs of the individual company.

When considering the use of the SFA scores a number of issues have to be solved. This report has shortly discusses possible ways of combining SFA and DEA scores, and ways of transforming the efficiency score into efficiency improvement target (X-factor). Our preliminary suggestion is that DEA and SFA scores could be combined by calculating an average of the scores. When setting the improvement targets, the differences in the time scale needed for removing inefficiency related to operational expenditure and depreciation have to be taken into account. On possible approach would be to assume that a given percentage of the inefficiency can be removed in a year and that operational expenditure and depreciation have different percentages. However, these questions have to be analysed thoroughly before the results are used in actual regulation.

As the results will be used for setting an improvement target for a regulatory period of four years, our suggestions is that the model is estimated once before the start of the period. These results would be used for setting the company specific improvement target. The results suggest that use more than one year data could be used to smooth out fluctuation when setting the improvement target for regulatory period of four years. If reliable data is available, a few year data could be used as a basis for setting the improvement targets. During the regulatory period, yearly results can be published for information purposes. This would give the companies information on their development in the same spirit as the DEA scores during the current regulatory period.

It has to be noted that the comprehensive data set used in the study was limited to one year, 2004. Designing the actual regulatory model would benefit from analysis of longer time period. Once the 2005 data set is available, it is recommended that it is analysed in detail.

## References


Afriat, S.N. (1972), Efficiency estimation of Production Functions, International Economics Review, 13, pp. 568-598.

Agrell, P. and P. Bogetoft, (2003a) *Dynamic Regulation, Report* AG2-V1, Norwegian Energy Directorate NVE, SUMICSID AB.

Agrell, P. and P. Bogetoft, (2003b) *Norm Models, Report* AG2-V2, Norwegian Energy Directorate NVE, SUMICSID AB.

Agrell, P. and P. Bogetoft, (2003c) *Use of Menu of Incentive Contracts, Report* AG2-V4, Norwegian Energy Directorate NVE, SUMICSID AB.

Agrell, P. and P. Bogetoft, (2003d) *Benchmarking for Regulation, Report* FP4, Norwegian Energy Directorate NVE, SUMICSID AB.

Agrell, P. and P. Bogetoft, (2003e) *Integrated, Parallel Energy Regulation, Report* FP5, Norwegian Energy Directorate NVE, SUMICSID AB.

Agrell, P. and P. Bogetoft, (2004) *NVE Network Cost Efficiency Model*, Norwegian Energy Directorate NVE, SUMICSID AB.

Agrell, P. J, P. Bogetoft, and J. Tind (2005), Scandinavian DEA Incentive Regulation in Electricity Distribution. *Journal of Productivity Analysiss*, 23, pp.173–201.

Aigner, D.J. and S.F. Chu (1968), On Estimating the Industry Production Function, *American Economic Review*, 58, pp. 826-839.

Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, Journal of Econometrics, 6, pp. 21-37.

Banker, R.D., *Hypothesis Test Using Data Envelopment Analysis*, Journal of Productivity Analysis, 7, pp. 139-159, 1996.

Battese, G.E. and T.J. Coelli (1992), Frontier Production Functions, Technical Efficiency and Panel Data: With Applications to Paddy Farmers in India, *Journal of Productivity Analysis*, 3, pp. 153-169.

Battese, G.E. and T.J. Coelli (1995), A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data, *Empirical Economics*, 20, pp. 325-332.

Bogetoft, P. (1994), Incentive Efficient Production Frontiers: An Agency Perspective on DEA, Management Science, 40, pp.959-968.

Bogetoft, P. (1997), DEA-Based Yardstick Competition: The Optimality of Best Practice Regulation, Annals of Operations Research, 73, pp. 277-298.

Bogetoft, P. (2000), DEA and Activity Planning under Asymmetric Information, *Journal of Productivity Analysis,* 13, pp. 7-48.

Bogetoft, P. and L. Otto, Notes to Production Economics, KVL. 2005.

Charnes, A., W. Cooper, A.Y. Lewin and L.M. Seiford (1994), *Data Envelopment Analysis: Theory, Methodology and Application*, Kluwer Academic Publishers.

Charnes, A., W.W. Cooper and E. Rhodes (1978) Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 2, pp. 429-444.

Charnes, A., W.W. Cooper and E. Rhodes (1979), Short Communication: Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 3, pp. 339.

Coelli, T., Rao, D.S.P. & Battese, G. (1998), *An Introduction to Efficiency and Productivity Analysi*s, Kluwer.

Cooper, W. W., Seiford, L. M. and Tone, K. (2000), *Data Envelopment Analysis*. Kluwer Academic Publishers.

Deprins, D., D. Simar, and H. Tulkens (1984) *Measuring Labor Efficiency in Post Offices*, pp. 243-267 in M. Marchand, P. Pestieau, and H. Tulkens, "The Performance of Public Enterprises: Concepts and Measurements", North Holland.

Edvardsen, D. F., and F. R. Førsund (2003), *International Benchmarking of Distribution Utilities.* Resource and energy Economics, 2003..

Edvardsen, D.F.(2004), Four essays on the measurement of productive efficiency, Ph.d. dissertation, Göteborg University.

Electricity Market Act (386/1995), Amendments until (1172/2004), Retrieved 20.03.2005 from http://www.energiamarkkinavirasto.fi/ (In Finnish and Swedish)

EMV (2004a) *Key figures for electricity network operations 2003*, Finnish Energy Market Authority, Retrieved 20.03.2005 from http://www.energiamarkkinavirasto.fi/ (In Finnish and Swedish)

EMV (2004b) *Guidelines for assessing reasonableness in pricing of electricity distribution network operations for 2005-2007*, Finnish Electricity Market Authority, 22 June 2004, Reg. no. 9/429/2004, Retrieved 20.03.2005 from http://www.energiamarkkinavirasto.fi/index.asp

EMV (2005) *Electricity market*, Finnish Energy Market Authority. Retrieved 20.03.2005 from http://www.energiamarkkinavirasto.fi/index.asp

Farrell, M.J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Series A, III, pp. 253-290.

Holmstrom, B. (1979), Moral Hazard and Observability, *The Bell Journal of Economics*, 10(1), 1979.

Holmstrom, B. (1982), Moral Hazard in Teams, *Bell Journal of Economics*, 13(2), 1982.

Honkapuro, Samuli; Kaisa Tahvanainen; Satu Viljainen; Jukka Lassila; Jarmo Partanen; Kimmo Kivikko; Antti Mäkinen ja Pertti Järventausta (2006) *DEA-mallilla suoritettavan tehokkuusmittauksen kehittäminen*, Energiamarkkinavirasto (In Finnish)

Kittelsen, S. A. C. (1993) *Stepwise DEA; Choosing Variables for Measuring Technical Efficiency in*

*Norwegian Electricity Distribution*, SNF-arbeidsnotat nr. A 55/93, SNF, Oslo.

Kittelsen, S. A. C. (1994) *Effektivitet og Regulering i Norsk Elektrisitetsdistribusjon*, SNF-rapport 3/94, SNF, Oslo.

Kittelsen, S. A. C. (1996) *DEA for NVE - Et Måleverktøy for Effektivitet i Elforsyningen*, SNF-rapport 85/96, SNF, Oslo.

Kittelsen, S. A. C. and Torgersen, A. M. (1993) *Teknisk Effektivitet i Norske Elektrisitetsfordelningsverk*, SNF-arbeidsnotat nr. A 27/93, SNF, Oslo.

Korhonen, P., Syrjänen, M., Tötterström, M. (2000) *Sähkönjakeluverkkotoiminnan kustannustehokkuuden mittaaminen*, Energiamarkkinaviraston julkaisuja 1/2000

Kumbhakar, S.C., S. Ghosh and J.T. McGukin (1991), A Generalize Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms, *Journal of Business and Economic Statistics*, 9, pp. 279-286.

Meeusen, W. and J. van den Broeck (1977), Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, *International Economic Review*, 18, pp. 435-444.

NVE (1997a) Retningslinjer for Inntektsrammen for Overføringstariffene. Report NVE, Norweigian Water Resources and Energy Administration, POB 5091, 0301 Oslo, Norway.

NVE (1997b) Benchmark. Publication 27/1997, Norwegian Water Resources and Energy Administration, POB 5091, 0301 Oslo, Norway.

Reifscheider, D. and R. Stevenson(1991), Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency, *International Economic Review*, 32, pp. 715-723.

Richmond, J. (1974), Estimating the Efficiency of Production, *International Economic Review*, 15, pp. 515-521.

Schmidt, P. (1976), On the Statistical Estimation of Parametric Frontier Production Functions, *Review of Economics and Statistics*, 58, pp. 238-239.

Simar, L. and P. Wilson (2000), A general methodology for bootstrapping in nonparametric frontier models, Journal of Applied Statistics, 27, 779-802.

Timmer, C.P. (1971), Using a Probabilistic Frontier Function to Measure Technical Efficiency, *Journal of Political Economy*, 79, pp. 776-794.

Tulkens, H. (1993), *On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit*, The Journal of Productivity Analysis, 4, 1993, pp. 183-210.