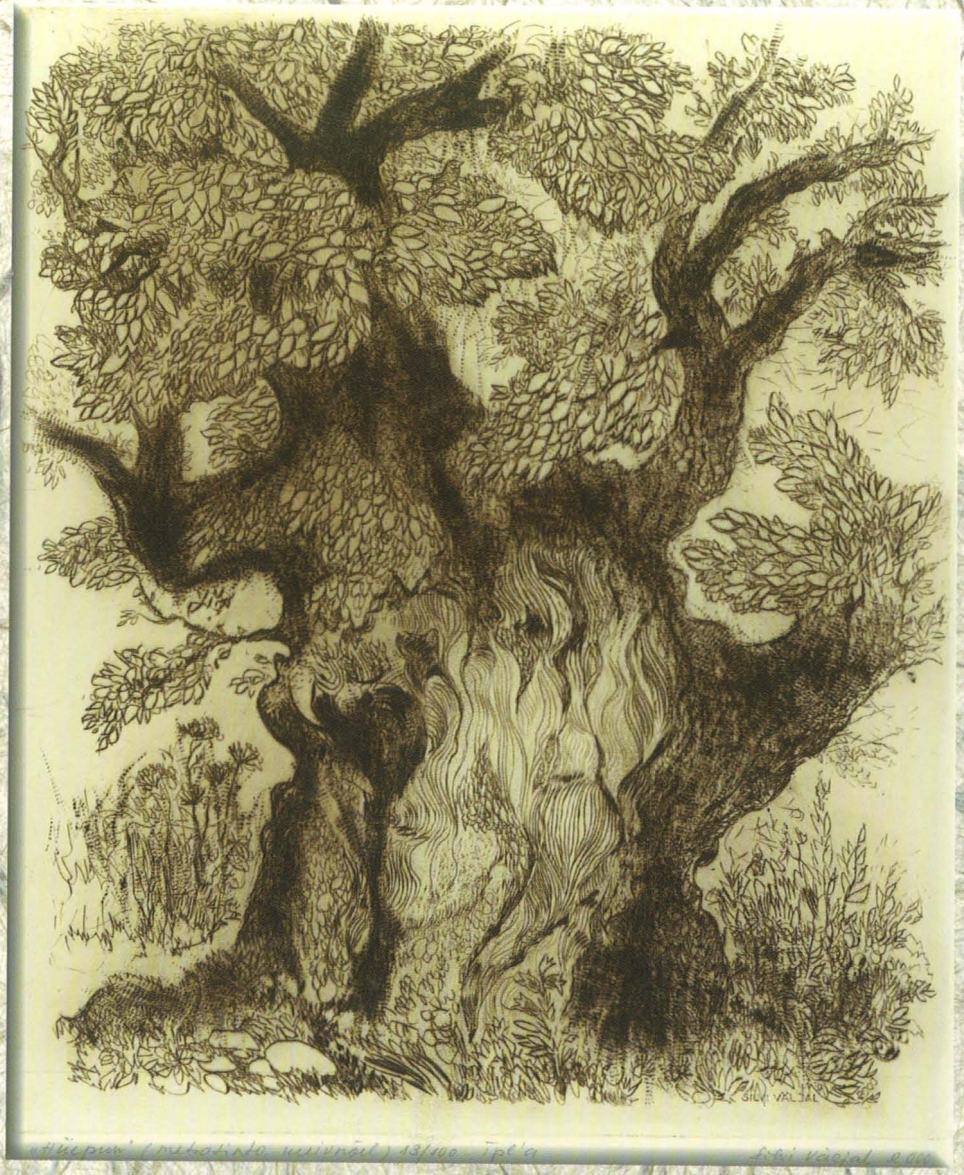
TALOUDEN TUTKIMUS JA PÄÄTÖKSENTEKO



KIRJOITUKSIA RAKENNEMUUTOKSESTA, KASVUSTA JA TALOUSPOLITIIKASTA

Toimittajat Kari Alho - Jukka Lassila - Pekka Ylä-Anttila

ETLA

ELINKEINOELÄMÄN TUTKIMUSLAITOS
The Research Institute of the Finnish Economy
Lönnrotinkatu 4 B 00120 Helsinki Finland

Sarja A 36 Series

Kari Alho – Jukka Lassila – Pekka Ylä-Anttila
(toimittajat, editors)

TALOUDEN TUTKIMUS JA PÄÄTÖKSENTEKO – Kirjoituksia rakennemuutoksesta, kasvusta ja talouspolitiikasta

ECONOMIC RESEARCH AND DECISION MAKING – Essays on Structural Change, Growth and Economic Policy

Juhlakirja Pentti Vartian 60-vuotispäivän kunniaksi
Essays in honour of Pentti Vartia on occasion of his 60th birthday

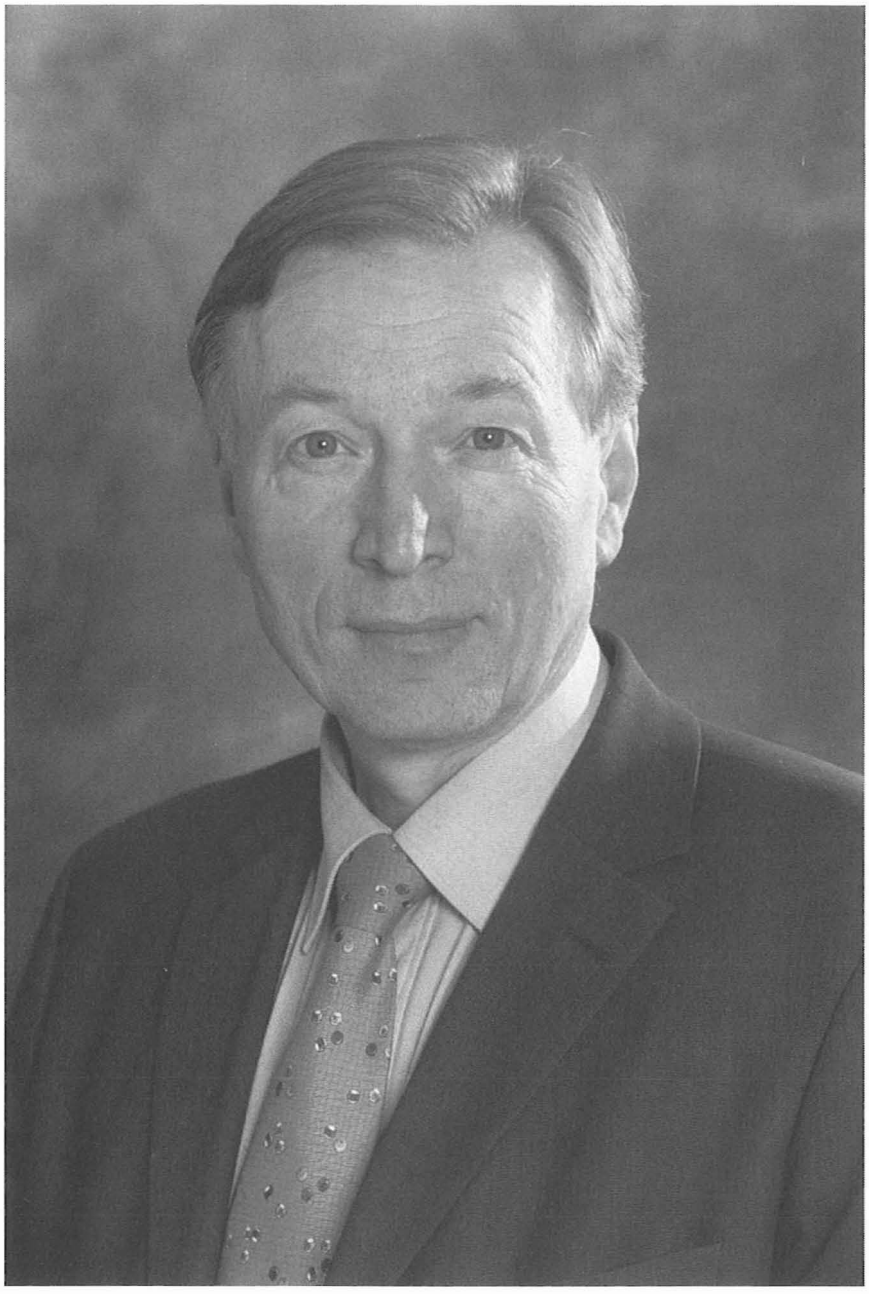
Julkaisija: Elinkeinoelämän tutkimuslaitos, ETLA
Kustantaja: Taloustieto Oy

Helsinki 2003

Kansi: MainosMayDay, Vantaa
Kannen kuva: Silvi Väljal "Hiiepuu"

ISBN 951-628-389-6
ISSN 0356-7435

Painopaikka: Yliopistopaino, Helsinki 2003



Pentti Vartia 60 vuotta 12.7.2003

SISÄLLYS

Saatesanat iii
Foreword iv

JOHDANTO

PENTTI VARTIA 60 VUOTTA – TUTKIJAN JA KESKUSTELIJAN MUOTOKUVA 1

OSA 1: KANSAINVÄLINEN TALOUS JA TALOUDEN KANSAINVÄLISTYMINEN

1. INTEGRAATIO, INSTITUUTIOT JA EU 9
   MIKA WIDGRÉN
2. EUROOPPA SIJOITTAJANA JA SIJOITUSKOHTEENA KANSAINVÄLISILLÄ PÄÄOMAMARKKINOILLA 27
   SINIKKA SALO
3. ECONOMIC INTEGRATION AND OPENNESS IN EUROPE AND EAST ASIA 45
   RAY BARRELL AND AMANDA CHOY
4. INTERNATIONAL BENCHMARKING AND ECONOMIC PERFORMANCE 69
   EINAR HOPE
5. FDI AND NORWEGIAN INDUSTRY 83
   PER HEUM

OSA 2: PITKÄ AIKAVÄLI – KASVUN LÄHTEET JA SELITYKSET

6. YRITYSKULTTUURIN ROOLI KAPITALISTISESSA TALOUDESSA 95
   VESA KANNIAINEN
7. THE VENTURE CAPITALIST AS A COMPETENT OUTSIDER 111
   GUNNAR ELIASSON
8. LASTENLASTEMME TALOUDELLISET MAHDOLLISUUDET 143
   JAAKKO KIANDER
9. OLEMMEKO MAAILMAN PARHAIMMISTOA? KANSALLISEN KILPAILUKYKYMME PUNTAROINTIA 163
   MATTI POHJOLA

10. Kasvun selitykset ja Suomen talouskasvu 183
Paavo Okko

11. Modelling Finnish Economic Growth: 1860-2001 199
Timo Teräsvirta and Dick van Dijk

OSA 3: TALOUSPOLITIIKKA JA TALOUDEN ENNUSTAMISEN VAIKEUS

12. Teollisuustuotannon käännepisteiden ennakointi barometri-aineiston avulla 223
Jussi Mustonen ja Penna Urrila

13. Analysing Short Heterogeneous Panel Data: The Demand for Labour in Finnish Firms 243
Markku Rahiala and Tapani Kovalainen

14. Models with Price Stickiness for Analysis of Monetary Policy 257
Seppo Honkapohja

15. Tuloerot ja tulonjako alueittain: Jakautuiko Suomi 1990-luvulla? 273
Heikki A. Loikkanen ja Risto Sullström

16. Devaluation Cycles in Sweden and Finland 287
Ulf Jakobsson

17. Talouspoliittinen keskustelu ja suuri lama 309
Jukka Pekkarinen

SAATESANAT

Tämä kirja on Pentti Vartian koti- ja ulkomaisten työtovereiden ja kollegoiden tapa muistaa 60-vuotiasta tutkijaa ja talouspoliittista keskustelijaa. Pentti on ETLAn – joka sekin on jo lähes kuudenkymmenen vuoden ikäinen – kautta aikojen pitkäaikaisin johtaja. Pentin aikana laitos on sekä kasvanut että muuttunut kansainväliseksi. Sen tutkimus- ja ennustetoiminta ovat saaneet vakiintuneen ja monella tavalla arvostetun aseman.

Tämän juhlakirjan kokoaminen ja toimittaminen on ollut mieluisa tehtävä. Ei ollut vaikeaa houkutella Pentin ystäviä ja kollegoita mukaan.

Kiitämme lämpimästi kaikkia kirjoittajia heidän työstään. Haluamme kiittää myös Mika Pajarista, Laila Riekkistä ja Petri Rouvista heidän panoksestaan kirjan toimittamisessa.

Helsingissä kesäkuussa 2003

Kari Alho Jukka Lassila Pekka Ylä-Anttila

FOREWORD

This book is a tribute to Pentti Vartia, a way for his Finnish and foreign colleagues to honour the 60th birthday of a highly regarded researcher and economic policy commentator. Pentti has served as the longest Director of ETLA, which too is almost 60 years old. During Pentti's time, the Institute has not only grown but has also internationalised itself in many respects. The Institute's research and forecasting activities have earned an established, and in many ways esteemed, position.

It has been a pleasure compiling this commemorative book and bringing it to completion. It was not hard to convince Pentti's friends and colleagues to collaborate on the project.

We extend our warm thanks to all the authors for their contributions. We also wish to thank Mika Pajarinen, Laila Riekkinen and Petri Rouvinen, who assisted in bringing this book to fruition.

Helsinki, June 2003

Kari Alho Jukka Lassila Pekka Ylä-Anttila

JOHDANTO

Pentti Vartia 60 vuotta – Tutkijan ja keskustelijan muotokuva

Pentti Lauri Ilmari Vartia syntyi Helsingissä 12. heinäkuuta 1943. Hän tuli ylioppilaaksi vuonna 1963. Opinnot johtivat aluksi hieman toiseen suuntaan kuin varsinainen elämänura eli Teknilliseen korkeakouluun, lentokoneen rakennuksen linjalle. Mutta samanaikaisesti virisi mielenkiinto talous- ja yhteiskuntatieteisiin. Pentti aloitti opintonsa myös Helsingin yliopiston valtiotieteellisessä tiedekunnassa.

Pentillä on siis kaksi akateemista tutkintoa: diplomi-insinöörin ja valtiotieteilijän. Jälkimmäinen ja kansantaloustiede nousivat jatko-opintojen keskeisiksi teemoiksi.

Pentti Vartia aloitti tutkijanuransa liittymällä uudelleen organisoidun ja laajenevan ETLAn palvelukseen vuonna 1971. Hän väitteli todella nopeasti tohtoriksi jo vuonna 1974 aiheenaan Suomen kansantalouden ekonometrinen malli.

Pentti nimitettiin kansantaloustieteen apulaisprofessoriksi vuonna 1975 Helsingin yliopistoon, mistä hän parin vuoden kuluttua siirtyi päätoimisesti ETLAn palvelukseen.

1970-luvun puolivälissä laitoksen vaiheikkaiden alkuvuosien jälkeen ETLAn toimitusjohtajaksi nimitettiin STK:n tilastojohtaja Tauno Ranta, jonka tehtäväksi tuli vakauttaa laitoksen toiminta. Hän valitsi aisaparikseen ja laitoksen tutkimusjohtajaksi Pentin vuonna 1977. Yhteistyö heidän välillään sujui mallikkaasti.

Tauno Rannan jäätyä eläkkeelle vuonna 1983 Pentti Vartia oli itsestään selvä valinta laitoksen toimitusjohtajaksi.

Kaksikymmentä vuotta on kulunut. ETLA on laajentunut ja kansainvälistynyt voimakkaasti. Laman aikana laitos joutui melkoiseen rahoituskriisiin, josta selvittiin kuitenkin hyvin suuntautumalla määrätietoisesti ulkopuolisen rahoituksen hankintaan ja käyttöön laitoksen tutkimuksessa.

Luottamustehtäviä aktiiviselle tutkijalle on kertynyt lukuisasti: esimerkiksi Yrjö Jahnssonin säätiössä, Suomen kulttuurirahastossa, Kansantaloudellisessa yhdistyksessä ja Taloustieteellisessä seurassa. Hänet on kutsuttu Suomen Tiedeseuran ja Teknillisten Tieteiden Akatemian jäseneksi. Pentti Vartia nimitettiin vuonna 2000 myös Elinkeinoelämän valtuuskunnan EVAn toimitusjohtajaksi. Tasavallan Presidentti myönsi Pentti Vartialle professorin arvonimen viime vuonna.

Pentti on vaikuttanut sekä puhumalla että kirjoittamalla. Hän on parhaimmillaan pohtiessaan vapaamuotoisesti yhteiskuntamme tilaa ja kehityssuuntaa, tiedeyhteisön keskustelua tai taloustieteen metodologisia pulmia. Selkeä luennoitsija on pitänyt talouspolitiikkaa koskevia luentoja Helsingin yliopistossa ja Helsingin kauppakorkeakoululla ja esiintynyt aktiivisesti yritysten ja järjestöjen tilaisuuksissa ja julkisessa sanassa.

Pentti on uransa aikana tarttunut moniin taloustieteen keskeisiin tutkimusalueisiin. Olemme valinneet niistä kolme – ehkä tärkeintä – , joiden mukaan tämä kirjakin on jäsennelty: kansainvälinen talous, talouden pitkän ajan kasvu sekä talouden ennustaminen ja talouspolitiikka.

KANSAINVÄLISEN TALOUDEN TUTKIJANA

Pentti Vartian tutkijauran keskeisenä johtoteemana on ollut Suomen talouden ja yhteiskunnan ongelmien tutkiminen osana kansainvälistä taloutta ja politiikkaa. Tämä on ohjannut hänen tutkimustoimintaansa kaikilla sen osa-alueilla ja läpi tutkijan uran vuosikymmenten.

On ollut luontevaa, että olemalla itse kansainvälisesti suuntautunut hän on myös johtanut ETLAn voimakkaaseen kansainvälistymiseen. Laitos on merkittävästi lähtenyt Suomen EU-jäsenyyden myötä mukaan myös EU-rahoitteisiin hankkeisiin.

Jo väitöskirjaa kirjoittaessaan 1970-luvun alussa Pentti lähti opin hakuun Hollantiin, maineikkaaseen Central Plan Bureau'hun, joka edustaa Jan Tinbergenistä juontuvaa hollantilaista ekonometrisen mallinrakennuksen traditiota. Sieltä löytyivät Suomen taloutta kuvaavan mallin ainesosat ja lähestymistavat. Väitöskirja synnytti pienen kiistan ekonometrikkojemme keskuuteen aiheesta "Tasot vai differenssit mallin muuttujien perusmuotona?".

Malliprojekti toimi myös sytykkeenä yhdelle ETLAn tieteellisesti merkittävimmistä löydöistä, nimittäin Pentin serkun, Yrjö Vartian indeksiteoreettisille pohdinnoille, jotka lähtivät liikkeelle lineaaristen tasesummien prosenttimuutosten analyysista. Indeksien konstruointi päätyi parhaimmillaan aina Econometrican sivuille.

Pentti Vartia vastasi mallin rakennuksen ohella ETLAn kansainvälisen talouden ennustamisesta. Työstä syntyi kirja raaka-aineiden hintojen ennustamisesta.

Pentti on osallistunut aktiivisesti AIECEn, Euroopan suhdanneinstituuttien liiton toimintaan ja on ollut sen puheenjohtaja vuodesta 1998 lähtien. Hän on ollut keskeisesti käynnistämässä ennustelaitosten EUROFRAME-

verkoston työtä. Nykyisin hän on myös Itävallassa toimivan systeemitutkimuksen instituutin, IIASAn, Suomen työryhmän puheenjohtaja.

Venäjä on ollut lähellä Pentin sydäntä ja tutkimusintressejä. Hän on ollut Suomen ja Venäjän välisen TT-yhteistyön yhteiskuntatieteellisen ryhmän puheenjohtaja. Hänen johdollaan ETLA on toteuttanut useita Venäjää koskevia hankkeita ja osallistunut mm. RECEP-tutkimuslaitoksen konsortioon Moskovassa, jossa laitoksen vastuulla oli Venäjän talousseuranta. Jo 1980-luvulla julkaistiin laitoksen toimesta idänkauppaa analysoinut perusteos "Venäjän-kauppa Suomen kansantaloudessa", jossa kriittisesti ja joitain hermostuttavastikin analysoitiin idänkaupan rakenteita.

ETLA on voimakkaasti vaikuttanut Suomen integraatiopolitiikan tutkimukseen. Suomi hakeutui kohti EY:tä 1990-luvun alussa ja ETLA ryhtyi aktiivisesti tutkimaan EY-jäsenyyden taloudellisia vaikutuksia sekä institutionaalisia kysymyksiä, kuten Suomen vaikutusvaltaa unionissa. Tästä lähti liikkeelle merkittävä tutkimus, jota nyt professoriksi edennyt Mika Widgrén edelleen jatkaa arvostettuna eurooppalaisena EU:n instituutioiden toiminnan ja reformin tutkijana. Edelleen ETLA on tutkinut aktiivisesti rahaliittoa ja viimeksi 2000-luvun puolella EU:n itälaajenemista, joka avaa uuden vaiheen Suomen ja Viron suhteisiin. Viro ja sen talous ovatkin olleet jo vuosikymmeniä lähellä Pentin sydäntä, ja hän rakensi yhteyksiä sinne jo 1970-luvulla.

Pentin monet luennot liittyvät Suomen asemaan maailmantaloudessa ja politiikkaan EU:ssa.

PITKÄ AIKAVÄLI – KASVUN LÄHTEET JA SELITYKSET

Taloustiede oli itse asiassa varsin pitkään vain vähän kiinnostunut pitkän aikavälin kasvusta. Tämä näkyi hyvin myös suomalaisessa tutkimuksessa. Pitkän aikavälin kehitysarvioita kyllä tehtiin jo 1960- ja 1970-luvuilla lähinnä hallinnon tarkoituksiin, mutta kiinnostus varsinaiseen kasvututkimukseen – kasvun selittämiseen ja ymmärtämiseen – syntyi vasta, kun uudet kasvuteoriat alkoivat 1980-luvulla kehittyä.

Samaan aikaan taloudellisia aikasarjoja tuottava taloushistoriallinen tutkimus edistyi monissa maissa. Laajoja kansallisia ja kansainvälisiä tietokantoja tuli tutkijoiden käyttöön. Empiirinen kasvun tutkimus eteni rinnan uusien kasvuteorioiden kehittymisen kanssa.

Pentti Vartian mielenkiinto pitkän aikavälin kasvuun tutkijana ja keskustelijana liittyy luontevasti tähän kehityskulkuun, mutta sillä on myös toinen ulottuvuutensa. Taloustutkijan näkökulma laajentui perinteisen taloustieteen ulkopuolelle: kulttuuriin, taiteeseen ja yhteiskunnan instituutioihin,

jotka ovat olleet viimeisten parinkymmenen vuoden aikana Pentin keskeisiä teemoja talouden kasvun ja muutoksen pohdinnassa. Pentin ajatusmaailmassa kansallinen kasvu- ja sivistysprojekti ovat saman asian kaksi puolta.

Vähintään yhtä tärkeää kuin pohtia taloudellisen kasvun lähteitä ja syitä, on pohtia kasvun vaikutuksia yhteiskunnan tulevalle kehitykselle. Miten käy eläke- ja muun hyvinvointijärjestelmän, jos jatkuvan kasvun kausi hiipuu? Miten voimme välttyä talouden kriiseiltä? Mitä tapahtuu, kun väestö vanhenee ja työvoima alkaa supistua? Näitä kysymyksiä Pentti on sekä pohtinut omissa kirjoituksissaan että edistänyt niitä koskevaa tutkimusta ponnekkaasti niin ETLAssa kuin muissakin yhteyksissä. Näitä kysymyksiä pohtii myös Pentin yhteistyökumppani Jaakko Kiander tämän kirjan artikkelissa. Hän päätyy - kuten Penttikin omissa pohdinnoissaan – verraten myönteiseen tulemaan: Suomi ei ole mitenkään väistämättä sidottu väestön vanhentumisen, supistuvan työvoiman tarjonnan ja hidastuvan talouskasvun kierteeseen.

Tulevaisuudella on tarjottavanaan aina useita vaihtoehtoja, joihin voidaan vaikuttaa – myös väestökehitykseen. Tämä on ollut yksi Pentin – useimmiten lopulta optimistiksi kääntyvän tutkijan – keskeisiä viestejä. Talouden ennustamisen vaikeus on tosiasia, mutta asiaa voi auttaa lisäämällä ymmärrystään tulevasta ja pyrkimällä vaikuttamaan hyvien vaihtoehtojen toteutumiseen.

Kiinnostus talouden pitkän aikavälin kehitykseen tarkoittaa tietysti paitsi tulevan kasvun ja kehityksen pohdintaa myös taloushistoriaan paneutumista. Tämä on ollut Pentin tutkijanuran luonteva vaihe ajattelun siirtyessä talouskehityksen suuriin linjoihin ja esimerkiksi talouden kriiseihin. Kriisit opettavat meitä aina palaamaan historiaan, sekä kansantalouksien että yritysten. Samat lainalaisuudet ovat toimineet aiemminkin, mutta ei koskaan täysin samalla tavalla.

Etiikan ja moraalin kysymykset ovat aina kiehtoneet Penttiä. Ne ovat luonteva osa talousjärjestelmien ja talouden pitkän aikavälin muutosten ja rakenteiden tutkimusta. Luottamus talouden instituutioihin ja niiden kyky muuttua – sosiaalinen pääoma – ovat juuri nyt ajankohtaisia kysymyksiä. Niin Pentillekin. Omalla tavallaan talouden instituutioihin ja organisoitumiseen tarttuu tässä kirjassa myös Pentin pitkäaikainen ystävä ja kollega Ruotsista, Gunnar Eliasson. Hänen artikkelinsa on samalla osoitus pohjoismaisesta yhteistyöstä, joka on ollut Pentin tutkijanuran keskeisiä elementtejä. Suuri osa tuosta yhteistyöstä on liittynyt juuri talouden rakenteiden ja pitkän aikavälin kasvun vertailevaan tutkimukseen.

TALOUSPOLITIIKKA JA TALOUDEN ENNUSTAMISEN VAIKEUS

Kun Pentti Vartia puhuu talouspolitiikkaan liittyvistä kysymyksistä, kuulija ei voi olla havaitsematta puhujan aitoa innostusta. Pentillä on selkeä käsitys sekä talouspolitiikan rajallisista mahdollisuuksista että näiden mahdollisuuksien käytön tärkeydestä. Keskeisiä toistuvia teemoja ovat kilpailukyky, kansantalouden sisäinen ja ulkoinen tasapaino, suhdanne- ja rakennepolitiikan tarve, tasapainoisesta kehityksestä huolehtiminen ongelmatilanteiden välttämiseksi, ja viime vuosina erityisesti taloudellisten kannustimien merkitys.

Pentti Vartia on ollut merkittävä pedagoginen hahmo taloudellisen ajattelun levittämisessä. Tämä koskee tiedotusvälineiden edustajien – Pentti on lähes aina mukana ETLAn Suhdanteiden ja tutkimusraporttien lehdistötilaisuuksissa – ja päättäjien lisäksi myös kollegoita. Hän on, kuten Jukka Pekkarinen kirjoituksessaan toteaa, pelkistäjä. Monimutkaisesta kokonaisuudesta pelkistyy olennainen asia kuvioksi tai tilastolliseksi riippuvuudeksi.

Talouspolitiikka tarvitsee ennusteita. Politiikan vaihtoehtoja on pohdittava sellaista kuvitteellista kehitystä vasten, jonka uskotaan toteutuvan ilman talouspolitiikan väliintuloa. Suomen talouden ennustaminen on ollut Pentin elämäntyötä runsaat kolme vuosikymmentä. Hänen näkemyksensä kiteytyy kirjassa "Talouden ennustamisen vaikeus".

ETLA on jo pitkään julkistanut tietoja ennustevirheistään suhdanneennusteen yhteydessä. Tämä on tapahtunut nimenomaan Pentti Vartian aloitteesta ja tahdosta. Tavoitteena on ollut antaa ennusteiden käyttäjille realistinen kuva ennusteiden tarkkuudesta tai tarkkuuden puutteesta. ETLAn ennusteet ovat aina olleet yhdistelmä kuvailevaa analyysia ja ennustelukuja. Jälkimmäisiä on ollut enemmän kuin useimmilla muilla suomalaisilla ennustelaitoksilla. Numeeristen ennusteiden hyvä puoli on se, että jälkikäteen voidaan saada määrällistä tietoa siitä, kuinka hyvin tai huonosti asiat ennalta arvioitiin. Laadullisesti oikeat ennusteet voivat olla käytännössä hyvin lähellä vääriä. Esimerkiksi 1980-luvun jälkipuoliskolla ennustejulkaisuissa nähtiin aivan oikein ongelmaksi talouden painottuminen liiaksi kotimaiseen kysyntään ja liian vähän vientiin. Tästä aiheutuvien vakavien vaikeuksien uhka nähtiin, mutta ennustenumeroiden ja toteutuneen kehityksen vertaaminen osoittaa, että 1990-luvun alun laman syvyys jäi täysin ennakoimatta.

Talouspoliittinen analyysi ja ennusteiden laatiminen vaativat tuekseen sekä teoreettista että soveltavaa kvantitatiivista tutkimusta. Pentti Vartia on koko työuransa ajan tehnyt, käynnistänyt ja kannustanut tällaista tutkimusta ja siihen olennaisena osana liittyvää tutkimusvälineiden rakentamista ja käyt-

töä. Empiiriset kokonaistaloudelliset mallit, teoreettiset päättelykehikot, tilastolliset analyysit eri tekijöiden vuorovaikutuksesta, aikasarja-analyysit ja kauas tulevaisuuteen kurkottavat simulointimallit ovat jokapäiväisessä käytössä siinä tutkimusyhteisössä, jossa Pentti Vartia toimii ja jota hän on keskeisesti itse ollut muovaamassa.

Tänä vuonna 60 vuotta täyttävä Pentti Vartia on edelleen niin nuorekas, että voimme rehellisesti todeta ETLAn, Etlatiedon ja Taloustiedon muodostaman ETLA-yhteisön olevan kovasti Pentin näköinen.

Kari Alho Jukka Lassila Pekka Ylä-Anttila

OSA 1:
KANSAINVÄLINEN TALOUS JA TALOUDEN KANSAINVÄLISTYMINEN

1. INTEGRAATIO, INSTITUUTIOT JA EU

Mika Widgrén*

MIHIN YLIKANSALLISIA INSTITUUTIOITA TARVITAAN?

Keskitetty ylikansallinen päätöksenteko on perusteltua, jos keskittämisellä voidaan hyödyntää yhteisistä päätöksistä seuraavia tehokkuus- ja mittakaavaetuja tai jos yhden jäsenmaan tai alueen päätökset heijastuvat myös muihin maihin tai alueisiin. Tehokkuusetuja voidaan saavuttaa muun muassa luomalla yhteisiä sääntöjä, standardeja tai hallinnollisia menettelyjä, jos nämä ovat aiemmin luoneet ainoastaan kustannuksia.

Kaupan esteillä on aina suora vaikutus kotimaisen tuotannon kanssa kilpailevan tuonnin kannalta. Esteiden poistamisella saavutetaan etuja, jos ulkomainen tuotanto on tehokkaampaa. Tämä voidaan rinnastaa niin sanottuun kaupan luomiseen. Myös erilaiset tuotestandardit tai säädökset käyvät esimerkiksi heijastusvaikutuksista erityisesti, jos kahden maan tai alueen sidos ulkomaankaupan välityksellä on tiivis. Erilaiset standardit ja säädökset vaativat, että vientiin ja kotimarkkinoille valmistetaan erilainen tuote, mikä saattaa luoda tehottomuutta. Tämän vuoksi ei olekaan yllättävää, että ylikansallisia jopa globaaleja instituutioita on syntynyt paljolti kansainvälisen kaupan ympärille.

On tunnettua, että maantieteellisesti rajoittunut kauppa-alue ei välttämättä pysty kohottamaan omaa taloudellista hyvinvointiaan huolimatta edellä kuvatuista tehokkuuseduista. Jos kauppa-alueen sidos muun maailman kanssa on tiivis, alueen sisäiset tehokkuusedut jäävät suhteellisesti ottaen vaatimattomiksi ja samalla muuta maailmaa kohti harjoitettu diskriminoiva politiikka luo tehottomuutta. Tämä vastaa niin sanottua kaupan syrjäytymistä. Se on sitä pienempi mitä tiiviimpi kauppasidos blokin jäsenmaiden välillä jo ennestään on. Tästä huolimatta regionalismin vaikutus muuhun maailmaan voi olla haitallinen, vaikka nämä haitat on mahdollista eliminoida niin sanotun Kemp-Wan -teoreeman perusteella. Sen toteutuminen ei kuitenkaan ole automaattista, mikä luo tarpeen globaaleille multilateraalisille instituutioille alueellisten kauppainstituutioiden vastapainoksi.

* Mika Widgrén on professori Turun kauppakorkeakoulussa. Kirjoittaja kiittää Kari Alhoa hyödyllisistä kommenteista.

Myös jäsenmaiden tai yhteistyöhön osallistuvien alueiden kansalaisten näkemykset vaikuttavat tavoiteltaviin yhteistyömuotoihin ja keskitetyn toimivallan laajuuteen. Jos näkemykset yhteisten päätösten hyödyistä ja haitoista poikkeavat kansalaisten keskuudessa maittain ja maiden sisällä paljon, päätösten keskittäminen tuottaa todennäköisemmin ristiriitoja voittajien ja häviäjien välille. Tällöin päätöksenteko on syytä säilyttää paikallisena. Mitä yhtenäisemmät näkemykset yhteistyön hyödystä ovat, sitä paremmin keskitetty päätöksenteko palvelee myös kansalaisia.

Kumpikin edellä mainituista edellytyksistä toteutuu todennäköisemmin toisiaan lähellä sijaitsevien maiden tai alueiden välillä. Läheisten alueiden luontainen sidos on vahvempi, alueellisten päätösten heijastusvaikutukset kohdistuvat todennäköisesti voimakkaimmin lähelle. Niin ikään pienehköillä yhtenäisillä alueilla kansalaiset tuntevat todennäköisemmin suurempaa tarvetta yhteistyöhön. Se, kuinka suurella voimalla etäisyys vaikuttaa tarpeeseen lisätä yhteistä päätöksentekoa, vaihtelee kuitenkin huomattavasti yhteistyömuodoittain. Tämän lisäksi globaalistuminen vähentää yleisesti etäisyyden merkitystä, joskaan ei – ainakaan vielä – ole poistanut sitä kokonaan. Etäisyyden merkitystä pitää yllä myös se, että tietyt ryhmäpäätöksiä tai yhteistä politiikkaa edellyttävät kysymykset ovat lähtökohtaisesti paikallisia.

Instituutioiden tehtävänä kansainvälisessä taloudessa on huolehtia niistä tehtävistä, joissa tietyt maat tai alueet tekevät yhteistyötä tai ovat integroituneet tai väljimmillään toimia jonkinasteisina koordinaattoreina. Hyvänä esimerkkinä ylikansallisesta itsenäisestä instituutiosta on Euroopan keskuspankki, väljää koordinaatiotehtävää hoitaa puolestaan esimerkiksi OECD. Kun kansainvälisen talouden instituutioilla on aidosti sitovaa ylikansallista tai ylialueellista päätös- ja toimivaltaa, myös niiden suunnittelun merkitys korostuu.

Kun yhteistyön tarve perustuu negatiivisten heijastusvaikutusten hallinnointiin, on periaatteessa kyse vastakkaisten intressien yhteensovittamisesta. Tämä ei kuitenkaan merkitse sitä, että kyseessä olisi nollasummapeli osapuolten kesken. Erityisesti yhdistettynä integraation mahdollisiin tehokkuusetuihin, kyse on lopulta pikemminkin molempia osapuolia hyödyttävästä asetelmasta. Esimerkkinä tästä voidaan pitää Euroopan integraation syntyvaiheita. Hiili- ja teräsyhteisön ja myöhemmin Euroopan talousyhteisön perustaminen pyrkivät ratkaisemaan sotateollisuuteen ja siihen liittyvään mahdollisen poliittisen ja sotilaallisen konfliktin uhan keinonaan aluksi hiili- ja terästuotteiden yhteismarkkinat ja talousyhteisön myötä vapaa kauppa ja tulliliitto, jotka ovat kiistatta lisänneet taloudellista hyvinvointia Euroopassa. Useat tutkimustulokset korostavat myös sitä, että tiivistyvä taloudellinen in-

tegraatio lisää turvallisuutta ja tätä kautta rohkaisee alueellisten toimivalta-alueiden syntyyn (Alesina & Spolaore 1997, Bolton & Roland 1997).

Institutionaalisesti ajatellen on merkittävää, että kun vuoden 1957 Rooman sopimuksella muodostettiin EEC-tulliliitto pelkän vapaakauppa-alueen sijaan, vaikutettiin merkittävästi talousyhteisön maista protektionistisimpien Italian ja Ranskan kauppapolitiikkaan, koska ne omaksuivat tullioliiton perustamisen myötä Benelux-maiden ja Saksan liberaalimman politiikan. Tästä huolimatta varsin tiiviin talousyhteisön muodostuminen koettiin uhkaksi Euroopan kauppasuhteille sekä myös laajemmin maailmankaupalle, mistä on osoituksena EFTAn synty vuonna 1959 ja GATTin Dillonin ja Kennedyn kierrokset vuosina 1959-61 ja 1962-64, jotka vauhdittivat globaalia vapaakauppaa Eurooppaan muodostuneiden alueellisten kauppablokkien vastapainoksi (esim. Sapir 1992). Kennedyn kierros myös pudotti EEC:n ulkotullitasoa huomattavasti – kenties tulliliitto-EEC oli helpompi neuvottelukumppani kuin se pelkäksi vapaakauppa-alueeksi institutionalisoituneena olisi ollut.

Tulliliiton muodostaminen oli merkittävä askel myös Euroopan integraation institutionalisoitumisessa. Vapaakauppasopimus tai -alue ei edellytä yhteisiä päätöksentekoinstituutioita, koska niihin osallistuvien maiden ulkoinen kauppapolitiikka säilyy autonomisena. Tulliliitto sen sijaan merkitsee yhteistä ulkoista kauppapolitiikkaa siihen osallistuvien maiden kesken ja rakentaa näin pohjan yhteisille ylikansallisille instituutioille ja päätöksenteolle.

Alueellisten integraatiosopimusten, kuten EU tai NAFTA, ja globaalin vapaakaupan välinen vuorovaikutus on jatkunut tämän jälkeenkin. Tämä kuvastaa Euroopan unionin merkitystä muun maailman kannalta, mikä puolestaan korostaa EU:n institutionaalisen rakenteen merkitystä. Samalla tuo vuorovaikutus on luonut oman instituutionsa WTO:n, jolla on keskeinen rooli toimia vahtikoirana alueelliselle integraatiolle tai suurten maiden kauppapolitiikalle. Taloustieteessä on harvoja tuloksia, joista valtaa yhtä suuri yksimielisyys kuin siitä, että globaali vapaakauppa lisää taloudellisista hyvinvointia. Vaikka WTO kykenee päättämään koko maailmankaupan säännöistä, perustuu sen päätöksenteko yksimielisyyteen ja jäsenmaiden hyväksyntään. Näin WTO:n rooli korostuukin sovittujen vapaakaupan sääntöjen valvojana.

Myös Euroopan taloudellinen integraatio rakentuu laajasti ymmärtäen vapaan kaupan pohjalle. Jos nykyisestä EU:n perustamissopimuksesta haluaa etsiä yhtä ydintä, sellaisen eittämättä muodostavat neljä vapautta: tavaroiden, palvelujen, työvoiman ja pääoman vapaa liikkuvuus. Kuinka avoin EU on sitten ollut WTO:n silmin? Tutkimustuloksissa vallitsee kohtalaisen laaja

yksimielisyys siitä, että keskimäärin varsin avoin. Selvänä poikkeuksena on EU:n maatalouspolitiikka, jonka siivittämä protektionismi on haitannut muuta maailmaa. Bourguignon ym. (2002) arvioivat, että EU:n yhteinen maatalouspolitiikka maksaa hyvinvointitappiona noin 300 euroa vuodessa jokaiselle EU-alueen kuluttajalle.

MITÄ EUROOPAN UNIONIN PITÄISI TEHDÄ?

Euroopan unionin toimivaltaa on vuosien varrella ulotettu alueille, joilla ei ole välttämättä kovin paljon tekemistä neljän vapauden – sisämarkkinoiden – kanssa. Näillä alueilla Euroopan laajuiset yhdentymisen hyödyt eivät ole yhtä kiistattomat, eikä yleinen mielipidekään yhtä kiistatta päätöksenteon keskittämisen kannalla. Tähän on osaltaan vaikuttanut unionin institutionaalisen rakenteen sekavuus.

EU:n toiminnan kehitystä ja laajenemista kuvaa esimerkiksi EU:n vuotuisen lainsäädännön määrä. Näin arvioiden unionin päätöksenteko on reilusti nelinkertaistunut 1970-luvun alkupuolelta vuoteen 2000.[1] Vuosina 1971-75 unioni teki noin 2500 säädöstä, direktiiviä tai päätöstä, vuosina 1996-2000 reilut 11 000.

EU-instituutio vaikuttaa monella tavalla jokaisen unionikansalaisen elämään. Nykyisellään se on kuitenkin kehittynyt osin hallitsemattomasti ja kaipaa selkeyttämistä, EU-instituution tarkempaa suunnittelua. Tämä edellyttää kolmen suuren valtakysymyksen ratkaisua. Ne ovat vallanjako unionin jäsenmaiden, vallanjako unionin toimielinten ja vallanjako jäsenmaiden ja unionin välillä. Kaikki kolme vaikuttavat siihen, minkälainen unioni pohjimmiltaan on ja mikä kansalaisten asema unionissa on.

Perustuslaki on keino selkeyttää vallanjakoa. Se saa myös unionikansalaisten laajan tuen. Keväällä 2003 julkaistun Eurobarometrin mukaan 65 prosenttia EU-kansalaisista kannatti perustuslakia, vain yhdeksän prosenttia vastusti sitä. Myönteisimmin perustuslakiin suhtautuvat italialaiset, kreikkalaiset, luxemburgilaiset, hollantilaiset ja ruotsalaiset, kaikki yli 70 prosentin kannatuksella. Penseimmin EU-perustuslakiin suhtautuvat puolestaan britit, tanskalaiset ja suomalaiset, mutta heidänkin keskuudessaan kannattajia on selvästi vastustajia enemmän.

Vallanjako jäsenmaiden ja unionin välillä on keskeinen kysymys käynnissä olevan liittovaltiokeskustelun kannalta. Euroopan unionissa on jo nyt pitkälle edenneitä piirteitä federalistisesta rakennelmasta sisämarkkinoiden, yhteisen valuutan ja rahapolitiikan kautta. Alesina ym. (2002) tarkastelevat, millä tasolla eri politiikkalohkoilla asioista tulisi päättää: paikallisella, mai-

den välisellä, kuten EU, vai globaalilla. EU-päätösten lukumäärän jyrkkä kasvu kertoo päätöksenteon keskittymisestä EU:ssa. Keskittymisen lisäksi päätöksenteko on levinnyt uusille politiikkalohkoille. Euroopan unionin yhtenä piirteenä on ollut se, että laajenemisten yhteydessä uusien jäsenmaiden lainsäädäntöä on harmonisoitu olemassa olevaan EU-lainsäädäntöön. Uusien jäsenmaiden eroavuudet ovat levittäneet unionipäätöksenteon toimialaa.

Vertailuksi Yhdysvaltojen laajetessa uusilta osavaltioilta ei edellytetty niiden lainsäädännön harmonisointia vanhojen jäsenosavaltioiden kanssa. Liittovaltio tarjosi vuoteen 1887 asti neljä vapautta sekä yhteisen kauppa-, ulko- ja turvallisuuspolitiikan. Osavaltioiden välisen taloudellisen kanssakäymisen säätely esimerkiksi verotusta tai kilpailuolosuhteita harmonisoimalla oli tätä ennen tuntematon käsite, eikä se ole sen jälkeenkään näytellyt samaa roolia kuin Euroopan integraatiossa (Rosenthal 1995).[2]

EU:ssa maatalouspolitiikan osuus kaikista päätöksistä on edelleen yli 40 prosenttia. Kaupan vapautuminen on vähentänyt tarvetta kaupan esteisiin liittyvään päätöksentekoon, mutta samalla sisämarkkinat ovat tuoneet kuvaan kilpailupolitiikan ja yrityslainsäädännön osaksi unionipäätöksentekoa. Kaupan ja laajasti ymmärrettyjen sisämarkkinoiden osuus unionipäätöksenteosta on noin kolmannes. (Alesina ym. 2002.)

Usein mainittu esimerkki politiikkalohkosta, jossa yhden maan tai alueen päätöksillä on heijastusvaikutuksia muihin maihin tai alueisiin, ovat ympäristökysymykset. Tämä puoltaa ylikansallista päätöksentekoa, mutta Euroopan unionin rooli ympäristöpolitiikassa on vähäinen. Myös kauppapolitiikan sekä pääoman ja työvoiman liikkumiseen liittyvillä päätöksillä on ulkoisvaikutuksia, jotka koskettavat muitakin alueita kuin sitä missä päätökset tehdään.

Sisämarkkinoilla on unionikansalaisten kiistaton tuki takanaan. Tätä voidaan pitää perusteena ottaa neljä vapautta osaksi EU-tason päätöksentekoa ja antaa niille samalla jopa perustuslaillinen asema.[3]

Kokonaan toinen kysymys on, kuinka paljon neljä vapautta heijastuu unionin päivittäiseen lainsäädäntöön ja unionitasolla tehtävien päätösten määrään muilla alueilla. Esimerkkinä toimivat yhteinen raha, kilpailupolitiikka ja yrityslainsäädäntö, joiden merkitys unionin päätöksenteossa on merkittävästi kasvanut sisämarkkinaohjelman myötä. Näkemykset näiden politiikkalohkojen oikeasta päätöksentekotasosta vaihtelevat.

Käytännössä yhteistä valuuttaa ei voida kirjata unionin perustuslakiin niin kauan kun osa jäsenmaista ei osallistu EMUun. Euron käyttöönoton vaikutus kansalaismielipiteeseen on ollut kuitenkin merkittävä. Kolme neljästä Euro-alueen kansalaisesta pitää EMUa hyvänä asiana. Käytäntö on nostanut

rahataloudellisen integraation suosiota myös Suomessa, jossa EMUn kannattajien osuus nousi euron käyttöönoton jälkeen 64 prosenttiin.

Alesina ym. (2002) ja Tabellini (2002) pitävät kilpailupolitiikkaa ja yrityslainsäädäntöä unionitason ja jopa globaalina WTO-tason asiana lähinnä ulkoisvaikutusten perusteella. Jos kansallisilla hallituksilla on mahdollisuus estää sisämarkkinoiden toimintaa suojelemalla kotimaisia yrityksiä, ne jossain yhteydessä myös näin usein tekevät. Unionitasolle keskitetyn päätöksenteon vastakohta on niin sanottu systeemikilpailu, jossa eri *toimivalta-alueet* (jurisdictions) voivat tarjota samanlaista julkishyödykettä ja unionikansalaiset voivat valita, mihin toimivalta-alueeseen kuuluvat. Ihmisten vähäinen liikkuvuus saattaa kuitenkin antaa paikallisille päätöksentekijöille monopolivaltaa, mikä rajoittaa systeemikilpailun toimintaa.

Sinn (2002) nostaa esille systeemikilpailun näkökulman, jossa eri maiden tai alueiden järjestelmät kilpailevat keskenään. Usein käytetty esimerkki systeemikilpailusta on verokilpailu. Esimerkkinä voidaan käyttää USA:ta, jossa kaikki välillinen verotus tapahtuu osavaltioissa tai vielä tätäkin paikallisemmin. Verokilpailusta kertoo esimerkiksi Oregonin liikevaihtoveron nollaverokanta, kun Kaliforniassa sovelletaan kuuden prosentin verokantaa. Osavaltioverokantojen lisäksi esimerkiksi Los Angeles county kantaa kahden prosentin alueellista liikevaihtoveroa. Välillinen verotus USA:ssa on esimerkki verokilpailusta, mutta myös hajautetusta päätöksenteosta. Sinn (2002) pitää mahdollisena, että systeemikilpailua voitaisiin soveltaa unionin kilpailupolitiikkaan sen sijaan, että se olisi unionitasolla tapahtuvan päätöksenteon asia.

Soveltaen yllä kuvattuja taloustieteellisiä periaatteita EU:n federalismiin Alesina ym. (2002) vetävät sen johtopäätöksen, että EU toimii liian vähän ympäristöpolitiikassa ja omissa ulkosuhteissaan (pl. kauppapolitiikka). Myös yhteiselle puolustus- ja turvallisuuspolitiikalle sekä rajat ylittävän rikollisuuden sekä terrorismin torjunnassa on taloustieteellisiä perusteita keskitetylle unionitason päätöksenteolle. Näissä kaikissa valtaosa unionikansalaisista on keskitetyn päätöksenteon kannalla, hajautetun päätöksenteon potentiaalisesti haitalliset ulkoisvaikutukset ovat suuret ja keskitetyn päätöksenteon mittakaavaedut merkittävät. Kansalaismielipiteen suhteen suomalaiset, ruotsalaiset ja britit muodostavat poikkeuksen. Kaikissa niissä suurempi osa kansasta vastustaa kuin puoltaa Eurobarometrin kyselyn perusteella ympäristökysymysten ja ulkosuhteiden ratkaisemista unionitasolla.

Maatalouspolitiikka on yksi keskeisimpiä unionitason asioita. Taloudellisten argumenttien pohjalta voidaan kuitenkin väittää, että kansallisen maatalouspolitiikan ulkoisvaikutukset ovat EU:n sisällä varsin pienet, jos

kansallisella maatalouspolitiikalla ei aleta tehdä aktiivista kauppapolitiikkaa. Nykyinen EU:n yhteinen maatalouspolitiikka on kuitenkin unionin selkeästi protektionistisin luomus. Kansalaismielipidekään ei pidä yhteistä maatalouspolitiikkaa kovin hyvänä asiana, koska se saa enemmistön tuen vain noin puolessa jäsenmaista. EU:n maatalouspolitiikan purkaminen on noussut esille itälaajenemisen yhteydessä – onhan se ollut myös laajenemista hidastava ja vaikeuttava kysymys.

EU:n yhteistä budjettia käytetään nykyisin lähes yksinomaan tulonjakoon ja vähäisessä määrin ulkosuhteisiin. Sen pääasialliset käyttökohteet ovat maatalouspolitiikka ja unionin köyhimpien alueiden tukeminen. EU:n maatalouspolitiikkaa tulee joka tapauksessa uudelleensuunnata suoran tulotuen suuntaan. Samalla huomattava osa sitä pitäisi kansallistaa edellyttäen kuitenkin, etteivät kansalliset ratkaisut luo kilpailun esteitä sisämarkkinoilla. EU:n yhteistä aluepolitiikkaa voidaan perustella alueiden epätasaisella taloudellisella kehityksellä tai lyhyemmällä aikavälillä niin sanotuilla epäsymmetrisillä taloudellisilla häiriöillä. Yleisesti nykyistä rakennerahastoin toteutettavaa mallia ei kuitenkaan voida pitää menestystarinana. Sekä alue- että maatalouspolitiikan kohdalla pätee myös, että päätöksentekijät EU-tasolla ovat jäsenmaita, mutta tuensaajat alueellisesti rajattuja. Tällöin tehokkaampi tapa organisoida nämä yhdessä perustuu könttäsummamaksuihin EU-budjetista jäsenmaille nykyisen koheesiorahaston tavoin jättäen tulonsiirtojen alueellinen kohdentaminen jäsenmaiden päätettäväksi.[4]

Yhteenvetona perustuslakiasiakirjaan kuuluvat luontevasti perus- ja ihmisoikeudet. Varsinaisen unionikompetenssin suhteen keskeiset elementit voidaan listata seuraavasti: neljä vapautta (sisämarkkinat), yhteinen raha, yhteinen budjetti, yhteiset ulkosuhteet myös kauppapolitiikkaa pidemmälle vietynä, sisäasiat siltä osin kun niillä on jäsenmaiden rajat ylittäviä vaikutuksia ja mahdollisuus erota unionista sekä joustava integraatio.

Edellä mainittujen mittakaavaetujen perusteella yhteinen ulko- ja turvallisuuspolitiikka, rajat ylittävän rikollisuuden torjunta, rajat ylittävät ympäristökysymykset ja unionin ulkorajan valvonta ovat kysymyksiä, jotka eivät nykyisellään ole unionin toimivallassa lainkaan tai hyvin heikosti. Niiden kuitenkin tulisi kuulua unionin toimivaltaan eli niin sanottuun ensimmäiseen pilariin. Myös näiden kysymysten nostamiselle unionitasolle on EU-kansalaisten enemmistön tuki.

Niin sanottu viides vapaus tarkoittaa äärimmilleen vietynä rajoittamatonta oikeutta perustaa toimivalta-alueita. Unionitasolla tämä tarkoittaa luontevasti joustavaa integraatiota tai, kuten Nizzan sopimus sitä kutsuu, vahvistettua yhteistyötä. Joustava integraatio sallii jäsenmaiden tai alueiden

muodostaa toimivalta-alueita, kunhan niiden toiminta ei ole ristiriidassa unionin yhteisen perustan kanssa. Keskeistä on, että nämä EU-integraatiota syvemmälle menevät yhteistyömuodot ovat avoimia. Komission rooli vahvistetun yhteistyön valvojana on keskeinen, mutta se ei saa muodostua esteeksi avoimien kumppanuuksien synnylle.

Eropykälän keskeinen merkitys on siinä, että se toimii yhtenä kontrollimekanismina missä tahansa ylikansallisessa instituutiossa. Jos integraatio etenee alueille, joita ei koeta hyödyllisiksi, jäsenmailla on mahdollisuus jättäytyä sen ulkopuolelle. On kuitenkin huomattava, että eropykälä tuskin voi koskea vain osaa jäsenyydestä. Eropykälän toinen merkitys on niissä kustannuksissa, joita jokainen ero aiheuttaa. Niiden minimoimiseksi on eduksi, jos instituution säännöt määrittelevät, mitä eron jälkeen tapahtuu ja missä mahdolliset erimielisyydet sovitaan.

KUINKA UNIONISSA TULISI PÄÄTTÄÄ?

Periaatteet

Ylikansalliset instituutiot saavat ajan myötä uusia jäseniä ja osa mahdollisesti jättää ne. Tämä heijastuu edelleen instituutioiden päätössääntöjen suunnitteluun. Myös niiden tulee perustua selkeisiin yleisesti hyväksyttyihin ja ennen kaikkea läpinäkyviin periaatteisiin. Itälaajenemisen ja vuodesta 1996 asti neuvotellun EU:n instituutioreformin välinen yhteys osoittaa, että unionissa asiat eivät olleet eivätkä vieläkään ole näin.

Institutionaaliset kysymykset, kuten äänimäärät neuvostossa ja maan paikkaluku Euroopan parlamentissa, ovat luoneet jokaisen laajenemisen yhteydessä hankalan neuvottelukysymyksen. Kun Suomi, Ruotsi ja Itävalta liittyivät unionin jäseniksi, neuvoston äänimääristä neuvoteltiin pitkään, koska erityisesti Ruotsi odotti saavansa viisi ääntä saamansa neljän sijaan.[5] Myös neuvoston enemmistösääntö herätti keskustelua, kun esimerkiksi Iso-Britannia halusi säilyttää sen 23 äänessä, mitä se EU12:ssa oli. Neuvoston kokonaismäärä nousi vuoden 1995 laajenemisessa 76:sta äänestä 87:ään. Britannian esitys olisi siis nostanut suhteellisen enemmistösäännön 71:stä noin 74:ään prosenttiin. Lopulta Ioninan kompromissilla päätettiin, että 23 äänen vähemmistöllä voidaan lykätä aloitteen lopullista hyväksyntää ja antaa mailе harkinta-aikaa muuttaa mieltään vähemmistön kannan mukaiseksi. Päätösten estämiseen tarvitaan kuitenkin 26 ääntä, mikä piti suhteellisen enemmistön ennallaan.

Instituutiot liittyvät hyvin suoraan myös itälaajenemiseen. Sen ennakkoehdoksi asetettiin jo vuonna 1996 EU:n instituutioiden uudistustyö, joka saatiin ratkaistuksi Nizzan huippukokouksessa joulukuussa 2000. Instituutiouudistus piti sisällään komission jäsenten lukumäärän rajaamisen, neuvoston äänimäärät, päätöksenteon tehokkuuden takaamisen myös laajentuneessa unionissa sekä säännöt niin sanotulle vahvistetulle yhteistyölle. Näistä komissiokysymyksen lopullinen ratkaisu jätettiin odottamaan itälaajenemisen yli ja neuvoston äänimäärien suhteen neuvoteltiin kompromissi, joka on saanut osakseen paljon kritiikkiä. Vaikka se muovasi äänten jakaumaa oikeaan suuntaan, tulos jäi puolitiehen. Samalla neuvoston suhteellista määräenemmistösääntöä nostettiin noin 71:stä prosentista 74:ään, mikä ei ainakaan paranna päätöksenteon sujuvuutta.

EU:n institutionaalisen rakenteen peruspuute on sen suunnittelun läpinäkymättömyys. Neuvoston äänimäärät ovat kuvaava esimerkki. Ne ovat kautta aikain määräytyneet niin, että maat on jaettu ryhmiin niiden koon mukaan ja kussakin ryhmässä maiden äänimäärät ovat samat. Ryhmien väliset rajat sekä ryhmien väliset äänimääräerot määräytyvät pitkälti mielivaltaisesti ja luovat ennen kaikkea tarpeen neuvotteluihin, kun unioni laajenee tai tapahtuu Saksojen yhdistymisen kaltainen muutos. Läpinäkyvyyden lisäksi on tavoiteltavaa, että instituutiot päivittävät näissä tapauksissa itse itsensä.

Euroopan unionin kompetenssin alueella voidaan ottaa esille kaksi mahdollista konstitutionaalista perusperiaatetta, joiden varaan päätöksenteko voidaan rakentaa:

1. Jokaisen EU-kansalaisen tulee olla samalla tavalla edustettuna hänen kotimaastaan riippumatta.
2. EU on kansalaisten ja kansojen unioni.

Näistä ensimmäinen on jokaisen demokraattisen päätöksentekojärjestelmän kulmakivi. Kaksitasoisessa päätöksentekojärjestelmässä EU-kansalaisten samanlaisen edustuksen toteuttaminen ei ole aivan suoraviivaista. Yleinen harhakäsitys on, että jäsenmaiden äänipainojen tulisi olla suoraan verrannollisia asukaslukuihin. Jos tavoitteena on, että kukin unionikansalainen on samalla tavalla edustettuna sen päätöksentekoelimissä, voidaan varsin helposti osoittaa, että suoraan asukasluvuista määräytyvä äänijako ei toimi. Tämä johtuu federaation kaksitasoisesta luonteesta: jäsenmaat valitsevat hallituksensa, jotka kokoontuvat ministerineuvostossa.[6]

Jos EU tekisi kaikki päätöksensä kansanäänestyksiin perustuen, olisi oikeudenmukaista antaa kullekin EU-kansalaiselle yksi ääni. Tällöin kullakin jäsenmaalla olisi sama määrä ääniä kuin asukkaitakin ja maiden ääniosuus

vastaisi tarkalleen niiden asukasluvun osuutta koko unionista. Tätä kutsutaan suhteelliseksi edustukseksi.

EU:n neuvoston päätökset tehdään kuitenkin käyttämällä edustukselista päätöksentekojärjestelmää. Jäsenmaat valitsevat ensin hallituksensa, joiden edustajat kokoontuvat tekemään EU-päätöksiä neuvostossa. Suhteellinen valtajako edellyttää, että kukin unionikansalainen edustaa vain itseään ja voi äänestää ikään kuin suoraan ministerineuvostossa. Kaksitasoisuus johtaa kuitenkin siihen, että näin ei ole.

Kansallisissa vaaleissa suurten maiden kansalaisilla on vähemmän valtaa kuin pienten maiden asukkailla omissa kansallisissa vaaleissaan, koska sekä pienissä että suurissa maissa kullakin äänioikeutetulla on yksi ääni, mutta suurissa maissa on enemmän äänioikeutettuja kuin pienissä maissa. Jotta kukin EU-kansalainen olisi samalla tavalla edustettuna neuvoston päätöksenteossa, pitää suurten maiden saada neuvoston äänestyksissä suurempi äänipaino kuin pienten maiden. Mutta kuinka paljon suurempi?

Lionel Penrose (1946) osoitti, että oikea kompensaatio perustuu maan asukasluvun neliöjuureen. Jotta kukin EU-kansalainen olisi yhtä vaikutusvaltainen neuvoston päätöksissä, tulisi maiden vaikutusvallan jakauman noudattaa niiden asukaslukujen neliöjuuriosuuksia.[7] Nizzan huippukokouksessa Ruotsin delegaatio esitti neliöjuurisäännön soveltamista päätettäessä ministerineuvoston äänimääristä.[8] Ehdotuksen hyvänä puolena on se, että EU:n neuvostossa se takaisi hyvällä tarkkuudella kaikkien unionikansalaisten tasapuolisen edustuksen neuvoston päätöksissä. Samalla se on erittäin läpinäkyvä systeemi.

Esimerkiksi Nizzan huippukokoukseen päättyneessä vuoden 2000 hallitusten välisessä konferenssissa tavoitteena oli ottaa huomioon EU:n luonne kansalaisten ja kansojen unionina. Tämä lienee ollut myös tärkein taustatekijä niin sanotulle *yksinkertaiselle kaksoisenemmistölle*, jota komissio ajoi omana ehdotuksenaan hallitusten välisessä konferenssissa ja Nizzassa.[9]

Yksinkertainen kaksoisenemmistö on standarditapa organisoida liittovaltion osavaltioiden välinen vallanjako. Se on siinä mielessä erikoinen äänestyssääntö, että se ei ole painotettu äänestys, mutta antaa silti eri maille ikään kuin erilaiset painot. Yksinkertainen kaksoisenemmistö edellyttää EU:n tapauksessa aloitteen taakse sekä jäsenmaiden enemmistön (nykyisin kahdeksan maata) sekä jäsenmaiden asukkaiden enemmistön (kun maat äänestävät käyttäen asukaslukuosuutta äänipainonaan).

Yllä mainittu periaate, jossa unioni ymmärretään jäsenmaiden ja kansalaisten unionina voidaan ymmärtää kahdella tavalla. Pohjimmiltaan se

määrittelee unionin luonteen federaationa, konfederaationa tai jäsenmaiden yhteisönä.

Yksinkertaisen kaksoisenemmistön sääntö perustuu ajatukseen, että vähemmistö maista tai asukkaista ei voi tehdä päätöksiä enemmistön kustannuksella. Näin sen taustalla piilee selkeä legitimiteetin ajatus. Myös yksinkertainen kaksoisenemmistö on erittäin läpinäkyvä periaate. Jos neuvosto samalla säilytetään keskeisimpänä päätöksentekijänä, annetaan ikään kuin lisäpaino hallitustenvälisyydelle, mutta nämä toteuttavat päätöksenteossaan federalistista mallia.

Liittovaltioiden päätöksenteossa, kuten esimerkiksi USA:n kongressissa, kaksoisenemmistön ajatus on rakennettu kaksikamarisen järjestelmän varaan. Osavaltiokomponentti on otettu huomioon senaatissa[10], jossa kullakin osavaltiolla on kaksi senaattoria. Valta on siis jaettu tasan osavaltioiden kesken. Asukaslukukomponentti on otettu huomioon edustajainhuoneessa, jonka paikat on jaettu osavaltioiden asukaslukuun perustuen. Molemmat kamarit ovat puolestaan samalla tavalla edustettuna, koska niiden molempien enemmistön hyväksyntä tarvitaan päätösten läpimenoon. Tällöin on kyse puhtaan federalistisesta mallista. Jos jäsenmaiden kamarille – EU:ssa neuvostolle tai USA:ssa senaatille – annetaan suurempi painoarvo on kyse konfederalistisesta mallista ääripäänään jäsenmaiden välinen yhteisö, jossa kansalaisten kamaria ei ole lainkaan.

Neuvoston äänimäärät

Kuviossa 1.1 sivulla 20 on tarkasteltu jokaisen EU-kansalaisen yhdenvertaisuuden toteutumista 27 jäsenmaan unionissa. Kuvio vertaa kahta erilaista tapaa allokoida äänet neuvostossa: Nizzan sopimuksessa määriteltyä tapaa ja menetelmää, jossa maiden asukasluvun neliöjuuri määrää äänimäärän, mutta aloitteen läpimeno edellyttää myös jäsenmaiden enemmistön taakseen. Jälkimmäinen piirre sisältyy myös Nizzan sopimuksessa määriteltyyn systeemiin. Molemmissa tapauksissa vaatimuksen tarkoitus on estää vähemmistöä jäsenmaista viemästä päätöksiä läpi.

Kuvio 1.1 vertaa Banzhaf-indeksillä mitatun vaikutusvallan jakauman neliöjuurisäännön välistä etäisyyttä maittaisten erojen neliösumman avulla. Normeerattu Banzhaf-indeksi voidaan kirjoittaa seuraavasti

$$\beta_i = \frac{s_i}{\sum_{i=1}^{n} s}$$

jossa s_i on niiden enemmistökoalitioiden lukumäärä, joissa maa i on mukana ja jotka se voi kääntää vähemmistöksi halutessaan. Normeerattu Banzhaf-indeksi mittaa siis äänestäjien vaa'ankieliasemien suhteellista osuutta. Neliöjuurisääntö toteutuu sitä paremmin mitä lähempänä Banzhaf-indeksit ovat maiden asukaslukujen neliöjuuriosuuksia

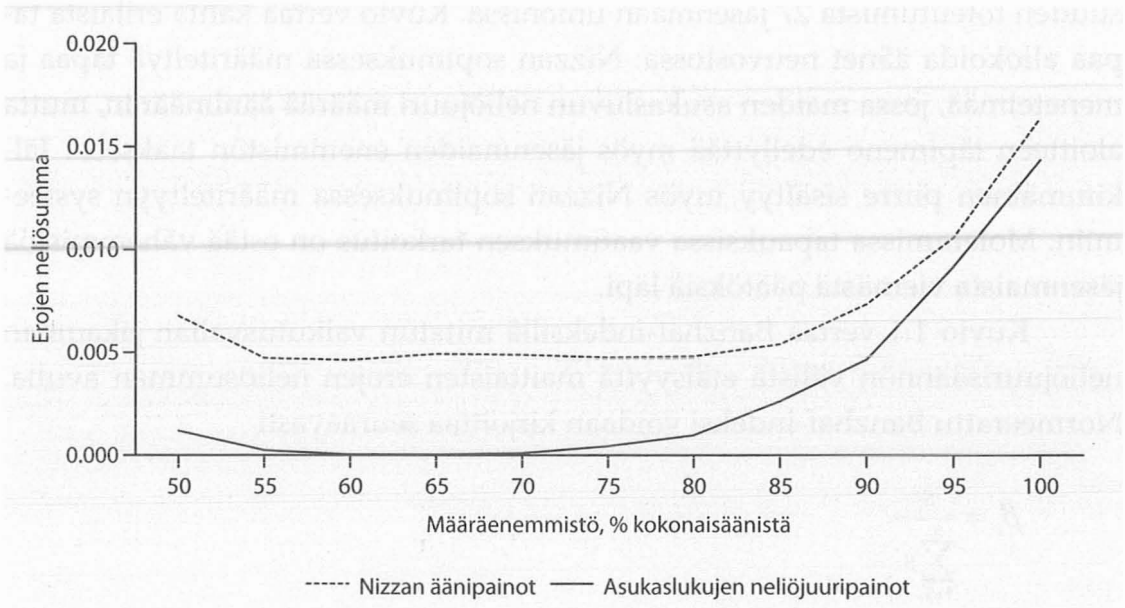
$$\beta_i^* = \frac{\sqrt{p_i}}{\sum_{j=1}^{n} \sqrt{p_j}}$$

jossa p_i on maan i asukasluku. Kuviossa 1.1 näiden kahden suureen etäisyyttä on mitattu erotusten neliösummalla $\sum_{i=1}^{n} \left(\beta_i^* - \beta_i\right)^2$.

Kuviosta 1.1 havaitaan, että molemmille ääntenmääräytymissäännöille on yhteistä niiden ja neliöjuurisäännön välisen eron vakaus varsin laajalla enemmistösääntöjen alueella noin 60 prosentista aina 80:een. Nizzassa sovittu äänijakauma tuottaa kuitenkin vaikutusvallanjakauman, joka poikkeaa jäsenmaiden asukaslukujen neliöjuuriin perustuvasta järjestelmästä. Kun vielä muistetaan neliöjuuripainotuksen läpinäkyvyys ja automaattisuus, on se ehdottomasti kannatettava tapa jakaa neuvoston äänet.

Myös EU-komission esittämä yksinkertainen kaksoisenemmistö on läpinäkyvä ja automaattisen tapa organisoida neuvoston päätöksenteko ja seuraa suoraan periaatteesta määritellä EU jäsenmaiden ja kansalaisten unioni-

Kuvio 1.1. Banzhaf-indeksin ja neliöjuurisäännön ero



na. Se ei kuitenkaan noudata neliöjuurisääntöä täsmällisesti. Tämä johtuu EU-maiden suhteellisen suurista kokoeroista. Suhteessa neliöjuurisääntöön yksinkertainen kaksoisenemmistö suosii systemaattisesti pieniä maita suurten kustannuksella.

Päätöksentekoelinten välinen valta

Ajatus Euroopan unionista kansojen ja kansalaisten unionina voidaan kiteyttää Euroopan parlamentin ja neuvoston välisiin voimasuhteisiin. Myös komission asema liittyy kiinteästi tähän, vaikka se voidaankin nähdä pikemminkin kysymyksenä Euroopan unionille keskitetyn ja jäsenmaille ja niiden kansalaisille hajautetun vallan suhteesta.

Myös päätöksentekoinstituutioiden välistä vallanjakoa voidaan arvioida kvantitatiivisesti. Klassiset valtaindeksit tai edes niiden instituutioiden välisen vallanjaon arviointiin tarkoitetut versiot eivät kuitenkaan sovellu tähän kovin hyvin, koska ne tarkastelevat päätöksentekoa simultaanisesti. Esimerkiksi Euroopan unionin päätöksentekoproseduurit ja erityisesti niihin osallistuvien tahojen siirtojärjestys on tarkoin määritelty, mikä luo oman lisämausteensa. Klassisten valtaindeksien logiikkaa voidaan kuitenkin soveltaa hieman muokattuna määrittelemällä pelaajien vaikutusvalta marginaalikontribuutiona äänestyspelin lopputulokseen eikä päätöksiä tekeviin koalitioihin, kuten esimerkiksi Banzhaf-indeksin tapauksessa edellä. Strateginen vaikutusvalta perustuu tällöin osittaisderivaattaan

$$\xi_i = \frac{\partial \chi(\lambda_i, \lambda_{-i})}{\partial \lambda_i}$$

jossa χ on pelin tasapaino funktiona pelaajan i preferensseistä λ_i ja muiden pelaajien preferensseistä λ_{-i} (ks. Napel & Widgrén 2002). Oletetaan, että mahdolliset pelin lopputulemat voidaan kuvata pisteinä yksikkövälillä, joka edustaa päätöksenteon politiikka-avaruutta ja jota voidaan tulkita esimerkiksi integraatio – ei-integraatiota -ulottuvuutena tai vasemmisto – oikeisto -ulottuvuutena. Preferenssit oletetaan tyypillisesti spatiaalisiksi, jolloin λ-termit edustavat pelaajien kannalta ideaalista politiikkalopputulosta ja tappiofunktiot ovat suoraan verrannollisia tehdyn politiikkapäätöksen ja pelaajan kannalta ideaalisen päätöksen väliseen etäisyyteen. Jos halutaan tarkastella apriorista vaikutusvaltaa, mikä on konstitutionaalisen analyysin kannalta keskeistä, voidaan esimerkiksi olettaa, että pelin status quo on normeerattu nollaan ja pelaajien ideaalipolitiikat ovat tasajakautuneita yksikkövälillä. Täl-

löin apriorinen strateginen vaikutusvalta on yllä olevan osittaisderivaatan odotusarvo.

Euroopan unionin keskeisin päätöksentekotapa on nykyisin niin sanottu yhteispäätösmenettely, joka luotiin Maastrichtin sopimuksessa 1992 ja uudistettiin Amsterdamin sopimuksella vuonna 1997. Menettelyn tavoitteena oli kohentaa Euroopan parlamentin asemaa. Aina vuoteen 1986 asti sillä ei ollut mitään formaalia roolia EU:n päätöksenteossa. Edelleenkin melko suuri osa päätöksistä tehdään vanhalla kuulemismenettelyllä, joka on puhtaasti hallitustenvälinen proseduuri.[11] Yhteispäätösmenettelyn nykyisessä versiossa komissio on aloitteentekijä, mutta elleivät neuvosto ja parlamentti hyväksy aloitetta, ne voivat hakea yhteistä kompromissia ilman komission väliin tuloa niin sanotussa sovittelukomiteassa. Näin komission rooli on yhteispäätösmenettelyssä rajattu niin sanottuun portinvartijaroolin. Toisin sanoen se voi olla tekemättä aloitteita, mutta ei aloitteen tehtyään enää estää parlamenttia ja neuvostoa poikkeamasta tästä. Tämä luo tehottomuutta, koska komissio voi olla tekemättä aloitteita tilanteissa, joissa yhteisistä päätöksistä olisi etua jäsenmaille sekä Euroopan parlamentille.

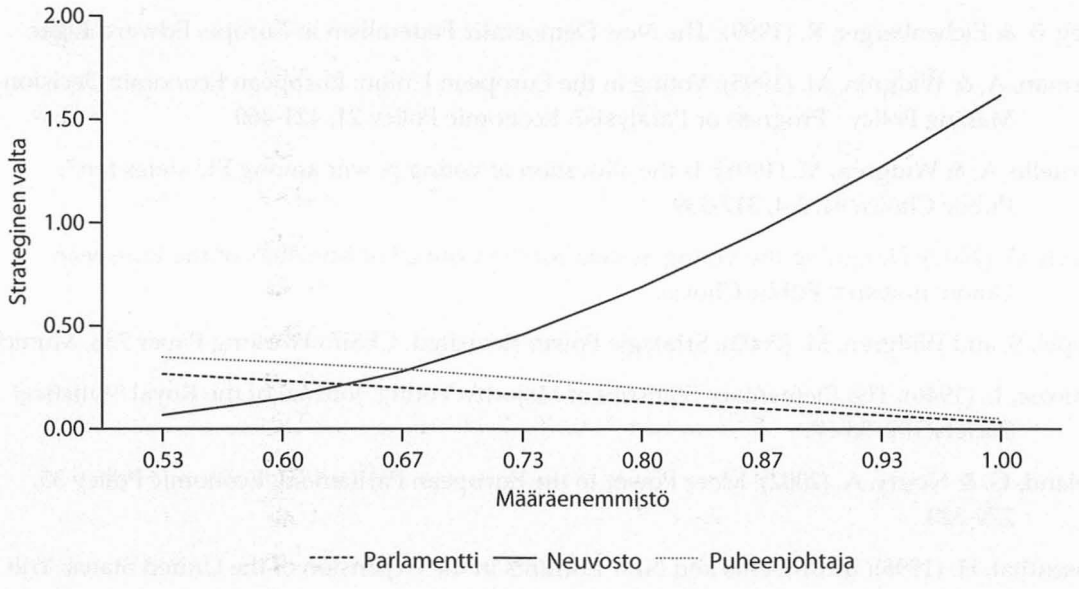
Kuviossa 1.2 on esitetty apriorinen strateginen vaikutusvalta yhteispäätösmenettelyssä, kun on oletettu, että neuvoston puheenjohtaja toimii sovittelukomitean puheenjohtajana. Kuvion vaaka-akselilla on neuvoston käyttämä enemmistösääntö, kun parlamentin on oletettu noudattavan yksinkertaista enemmistöä. Pystyakseli kuvaa strategista valtaa, joka vaihtelee nollan ja kakkosen välillä (ks. Napel & Widgrén 2002, Widgrén 2003).

Kuviosta 1.2 havaitaan, että neuvosto ja parlamentti ovat yhtä vaikutusvaltaisia, kun neuvoston määräenemmistö on hieman yli 60 prosenttia ja neuvoston ja sen puheenjohtajan valta yhtäläistyvät kahden kolmasosan enemmistön kohdalla. Jos neuvoston puheenjohtaja on yksi jäsenmaista, kuten nykyisin, parlamentti omaa kaikilla neuvoston määräenemmistösäännöillä selvästi vähemmän strategista valtaa kuin koko neuvosto. Jos nykyisessä yhteispäätösmenettelyssä halutaan pyrkiä noudattamaan ajatusta unionista kansojen ja kansalaisten unionina, neuvoston puheenjohtajan (presidentin) tulisi olla jäsenmaista ja neuvostosta riippumaton taho. Konventissakin keskusteltu kahden hatun presidentin malli on tällöin looginen vaihtoehto, koska siinä komission puheenjohtaja johtaa myös neuvostoa. Komissio saa tässä mallissa nykyistä vahvemman roolin, millä on se hyvä puoli, että kansalaisten ja kansojen unionin ajatus voidaan toteuttaa ja samalla komission aktiivisempi aloitteentekijärooli poistaa edellä mainitun tehottomuusongelman. Potentiaalisena haittapuolena voidaan nähdä komission vahvistumisesta ja aloitteenteon monopolista juontava päätöksenteon keskittyminen.

Jos komission aseman korostuminen koetaan ongelmana, on yhteispäätösmenettelyyn mahdollista lisätä kilpailullisia elementtejä, joilla komission aloitteenteon monopolia voidaan purkaa. Jos Euroopan parlamentille annetaan komissiosta riippumatonta aloitteentekovaltaa sovittelukomiteassa, kuviosta 1.2 havaitaan, että suunnilleen nykyisellä 71 prosentin enemmistösäännöllä neuvoston ja parlamentin sekä sovittelukomitean aloitteentekijän yhteenlaskettu valta ovat yhtä suuret. Tämä vastaa tilannetta, jossa neuvosto ja parlamentti voivat päättää komission aloitteesta poiketen, mutta se edellyttää neuvoston laajemman hyväksynnän. Samalla tämäntyyppinen ratkaisu vaatisi erillisen aloitteentekoinstituution parlamentin sisälle.

Yhteenvetona yhteispäätösmenettelyn heikkoutena voidaan nykyisin pitää sen potentiaalista tehottomuutta sekä parlamentin ja neuvoston aseman epäsuhtaa, joka johtuu neuvoston käyttämästä korkeammasta määräenemmistöstä sekä puheenjohtajamaan korostuneesta asemasta. Komission roolin vahvistaminen ratkaisee periaatteessa nämä molemmat, mutta päätöksenteon keskittymisen pelko edellyttää, että parlamentilla ja neuvostolla on mahdollisuus poiketa komission aloitteesta vaikeutetulla menettelyllä. Jos parlamentille annetaan tässä aloitteenteko-oikeus, neuvoston määräenemmistö olisi korkeampi kuin normaalimenettelyssä. Vähemmän institutionalisoidussa mallissa normaalimenettelyä korkeampi enemmistö sekä parlamentissa että neuvostossa voisivat sopia yhteisestä kompromissista.

Kuvio 1.2. Strateginen valta yhteispäätösmenettelyssä



KIRJALLISUUTTA

Alesina, A. & Spolaore, E. (1997): On the Number and Size of Nations, Quarterly Journal of Economics 112, 1027-1056.

Alesina, A., Angeloni, I. & Schuknecht, L. (2002): What Does the European Union Do?, CEPR Discussion Paper 3115.

Baldwin, R., Berglöf, E., Giavazzi, F. & Widgrén, M. (2000): EU Reforms for Tomorrow´s Europe, CEPR Discussion Papers 2623.

Baldwin, R., Berglöf, E., Giavazzi, F. & Widgrén, M. (2001): Nice Try: Should the Treaty of Nice Be Ratified?, Monitoring European Integration 11, Centre for Economic Policy Research, London.

Bolton, P. & Roland, G. (1997): The Breakup of Nations: A Political Economy Analysis, Quarterly Journal of Economics 112, 1057-1090.

Buchanan, J. (1991): An American Perspective on Europe's Constitutional Opportunity, Cato Journal 10, 619-629.

Casella, A. & Frey, B. (1992): Federalism and Clubs: Towards an Economic Theory of Overlapping Political Jurisdictions, European Economic Review 36, 639-646.

Dewatripont, M. & Giavazzi, F. & von Hagen, J. & Harden, I. & Persson, T. & Roland, G. & Rosenthal, H. & Sapir, A. & Tabellini, G. (1995): Flexible Integration, Centre for Economic Policy Research, London.

Felsenthal, D. and Machover, M. (2001): The Treaty of Nice and Qualified Majority Voting. Social Choice and Welfare 18, 431-64.

Frey, B. & Eichenberger, R. (1999): The New Democratic Federalism in Europe, Edward Elgar.

Kirman, A. & Widgrén, M. (1995): Voting in the European Union: European Economic Decision-Making Policy : Progress or Paralysis?, Economic Policy 21, 421-460.

Laruelle, A. & Widgrén, M. (1998): Is the allocation of voting power among EU states fair?, Public Choice 94: 3-4, 317-339.

Leech, D. (2002): Designing the Voting System for the Council of Ministers of the European Union, ilmestyy Public Choice.

Napel, S. and Widgrén, M. (2002): Strategic Power Revisited. CESifo Working Paper 736, Munich.

Penrose, L. (1946): The Elementary Statistics of Majority Voting. Journal of the Royal Statistical Society 109, 53-57.

Roland, G. & Noury, A. (2002): More Power to the European Parliament, Economic Policy 35, 279-320.

Rosenthal, H. (1995): Incumbents and New Entrants in the Expansion of the United States, Yrjö Jahnsson European Integration Lectures 10.

Sinn, H-W. (2002): The New Systems Competition, Yrjö Jahnsson Lectures, Blackwell.

Tabellini, G. (2002): Principles of Policymaking in the European Union: an Economic Perspective, CESifo Forum Vol. 3, No. 2, 16-22.

Widgrén, M. (1994): Voting Power in the EU and the Consequences of Two Different Enlargements. European Economic Review} 38, 1153-70.

Widgrén, M. (2001): Euroopan integration talous ja politiikka, Taloustieto.

Widgrén, M. (2003): Enlargements and the Principles of Designing EU Decision-Making Procedures, CESifo Working Paper 903, Munich.

LOPPUVIITTEET

[1] Yksityiskohtaisemmasta kuvauksesta ja päätöksenteon määrästä eri politiikkalohkoilla, ks. Alesina et al. (2002).

[2] Yhteinen valuutta implementoitiin Yhdysvalloissa sisällissodan aikana asettamalla vero muiden valuuttojen kuin dollarin käytölle. Huomattakoon, että Yhdysvalloissa yhteisen valuutan käyttöön otto vei noin 90 vuotta, EU:ssa 55.

[3] Ks. tarkemmin myös Alesina ym. (2002) ja Dewatripont ym. (1995). Neljän vapauden asemasta Euroopan taloudellisen integraation ytimenä ollaan yleisesti varsin yksimielisiä. Mutkikkaampi kysymys on jo, miten neljän vapauden tehokkaan implementoinnin tulisi heijastua unionin päätöksentekoon.

[4] Tästä argumentista, ks. Tabellini (2002).

[5] Analyysista esimerkiksi Widgrén (1994).

[6] Tarkkaan ottaen jäsenmaiden kansalaiset eivät valitse hallitusta suoraan, ellei käytössä ole kaksipuoluejärjestelmä. Muuten asetelma on kolmivaiheinen. Tarkastelu lähteekin siitä, että kaikilla jäsenmaiden kansalaisilla on yhtä suuret vaikutusmahdollisuudet vaikuttaa hallituksen kokoonpanoon. Tämä nostaa esille kuitenkin myös tärkeän kysymyksen siitä, pitäisikö ministerineuvoston jäsenet valita suoralla vaalilla, kuten senaattorit USA:ssa.

[7] Lisäongelmana on, että äänipainot ja vaikutusvalta eivät ole tyypillisesti yhtä suuria. Äänet ovat vain resurssi, jolla valtaa käytetään. Klassinen esimerkki on tilanne, jossa kolmesta äänestäjästä kahdella on 49 ääntä ja kolmannella kaksi. Jos päätökset tehdään yksinkertaisella enemmistöllä, on helppo nähdä, että äänestäjät ovat yhtä vaikutusvaltaisia, koska kukin on vaa'ankielenä kahdessa enemmistökoalitiossa, vaikka äänimäärien perusteella tilanne näyttää hyvin epäsymmetriseltä. Tämä tekee painotetun äänestyksen suunnittelun käytännössä hankalaksi. Euroopan unionissa erot vaikutusvallan ja ääniosuuksien välillä ovat kuitenkin kohtalaisen pienet ja pienenevät itälaajenemisen jälkeen (Widgrén 2003).

[8] Huomattakoon kuitenkin, että Ruotsin delegaatio ei tuntenut Penrosen tulosta. Ehdotuksessa laskettiin kunkin maan asukasluvun neliöjuuren osuus unionin jäsenmaiden asukaslukujen neliöjuurten summasta prosentteina ja tämä kerrottiin kahdella. Pyöristys lähimpään kokonaislukuun antoi sen jälkeen maan äänimäärän.

[9] Nizzan huippukokouksen alla harkittiin vielä viittä eri vaihtoehtoa neuvoston päätöksenteon uudistamiseksi. Näistä mikään ei sellaisenaan voittanut. Vaihtoehtojen yksityiskohtaisesta vertailusta, ks. Baldwin ym. (2000) ja Nizzan tulosten arviosta, ks. Baldwin ym. (2001), Felsenthal & Machover (2001) tai Leech (2002).

[10] USA:n laajenemisen yhteydessä ei missään vaiheessa asetettu kyseenalaiseksi sitä, pitääkö pienille osavaltioille antaa senaatissa sama edustus kuin osavaltioille, joissa asukkaita on enemmän (ks. Rosenthal 1995).

[11] Kuvauksesta päätöksenteon yksityiskohdista, ks. Widgrén (2001).

2. EUROOPPA SIJOITTAJANA JA SIJOITUSKOHTEENA KANSAINVÄLISILLÄ PÄÄOMAMARKKINOILLA

Sinikka Salo*

Euroopan rahaliiton ja EU:n sisämarkkinaohjelman ajankohtaisuuden vuoksi on luonnollista, että Euroopassa ja varsinkin euroalueella on viime vuosina kiinnitetty paljon huomiota alueen sisäisten rahoitusmarkkinoiden toimintaan, tehostamiseen ja yhtenäistämiseen. Euroopan laajuisten tehokkaiden rahoitusmarkkinoiden muodostuminen vasta mahdollistaa sisämarkkinoiden täysimääräisen hyödyntämisen. Tämä Euroopan sisäinen näkökulma ei kuitenkaan ole yksin riittävä, kun arvioidaan markkinoiden toimintaa ja kehittämistarpeita. Euroopan rahoitusmarkkinat ovat myös osa globaaleja, kansainvälisiä rahoitusmarkkinoita ja yksi niiden kaikkein tärkeimmistä tehtävistä on liittää euroalueen talous muuhun maailmantalouteen. Se, miten hyvin ne tässä tehtävässä onnistuvat, vaikuttaa paitsi euroalueen taloudelliseen menestykseen, myös maailmantalouteen kokonaisuudessaan.

Seuraavassa luon tiiviin katsauksen euroalueen kansainväliseen rahoitusmarkkinarooliin ja pyrin paikallistamaan muutamia ongelmakohtia tämän, Euroopan kannalta keskeisen roolin kehittämisessä. Aloitan vertailemalla euroalueen rahoitusmarkkinoita Yhdysvaltain ja Japanin rahoitusmarkkinoihin. Tämän jälkeen tarkastelen sitä, miten kansainvälisten rahoitusmarkkinoiden eri tehtävät sivuavat Euroopan talouspolitiikan nykyisiä ongelmia. Lopuksi otan esiin eräitä tulevan kehityksen kannalta keskeisiä kysymyksiä.

EUROOPAN RAHOITUSMARKKINAT GLOBAALISSA VERTAILUSSA

Talous- ja rahaliiton perustamisen myötä Eurooppaan syntyi maailman toiseksi suurin valuutta-alue. Bruttokansantuotteella mitaten sen kansantalous on noin kolme neljäsosaa Yhdysvaltain kansantaloudesta ja noin puolitoista-

* Sinikka Salo on Suomen Pankin johtokunnan jäsen.

kertainen Japanin kansantalouteen verrattuna. Kansainvälisessä kaupassa euroalueen suhteellinen merkitys on vielä suurempi kuin BKT-vertailujen perusteella voisi odottaa, koska euroalue on kauppavirtojen valossa Yhdysvaltoja jonkin verran avoimempi talousalue, mitattuna esimerkiksi tavaroiden ja palvelusten ulkomaankaupan BKT-suhteella. Viime vuosina euroalueen tavaroiden ja palvelujen vienti on ollut absoluuttisesti jopa jonkin verran suurempi kuin Yhdysvaltojen.

Myös rahoitusmarkkinoilla euroalue on erittäin merkittävä tekijä. Euroalue on Yhdysvaltain jälkeen maailman toiseksi tärkein säästäjä ja myös toiseksi tärkein investoija. Euroopan rahoitusmarkkinat ovat kooltaan maailman toiseksi suurimmat. Erityisesti pankkijärjestelmän koon suhteen Euroalue kilpailee tasavertaisesti Yhdysvaltojen kanssa, mikä näkyy esimerkiksi pankkien luotonantoa kuvaavissa tilastoissa. Maailman kolmannen rahoitusmahdin Japanin rahoitusmarkkinat ovat useimpien mittarien mukaan selvästi euroalueen markkinoita pienemmät.

Euroopan ja Yhdysvaltain rahoitusmarkkinoiden kokovertailuja vaikeuttaa tunnetusti se, että näiden alueiden rahoitusjärjestelmät ovat rakenteeltaan erilaiset. Yhdysvaltojen rahoitusmarkkinat perustuvat paljon enemmän julkisesti noteerattaviin arvopapereihin kuin mannereurooppalainen rahoitusjärjestelmä, joka on perinteisesti ollut pankkikeskeinen. Niinpä Yhdysvaltojen osakemarkkinat ovatkin kooltaan moninkertaiset Euroalueen osakemarkkinoihin verrattuna. Sama pätee julkisesti noteerattujen velkapaperien määrään.

Arvopaperimarkkinoihin perustuva vertailu antaa kuitenkin vääristyneen ja liian vaatimattoman kuvan euroalueen rahoitusresursseista. Todellisuudessa alueen rahoitusresurssien määrä riippuu enemmän alueella tapahtuvasta säästämisestä kuin siitä, kuinka suuri osa varallisuudesta on markkinoilla noteerattua. Arvopaperimarkkinoiden toiminta ei silti ole vailla merkitystä. Julkisten arvopaperimarkkinoiden koko kertoo markkinoiden kyvystä

Taulukko 2.1. Pankkien luotonanto eri talousalueilla

	Vuoden 2001 lopussa, mrd. euroa	
	Kotimainen	Alueen ulkopuolelle
USA	10175	558
Euroalue	9311	2407
Japani	6455	820

Lähteenä on International Financial Statistics.

koota ja käsitellä sijoittajien kannalta olennaista informaatiota, ja näiden markkinoiden pienuus taas antaa viitteitä informaatioperustaan ja infrastruktuuriin liittyvistä kehitystarpeista. Niinpä on ilmeistä mm. se, että euroalueen markkinoiden hajanaisuus ja siihen liittyvät korkeat arvopaperikaupan kustannukset haittaavat markkinoiden toimintaa ja ovat siten hidastaneet niiden kasvua. Muita, täydentäviä selityksiä euroalueen arvopaperimarkkinoiden suhteelliselle pienuudelle voi ilmeisesti löytää eläkejärjestelmien ja asuntorahoitusjärjestelmien eroista Euroopan, Yhdysvaltain ja Japanin välillä.

On mielenkiintoista todeta, että vaikka Yhdysvaltain pörssinoteerattujen yritysten yhteenlaskettu markkina-arvo onkin moninkertainen verrattuna euroalueen yritysten markkina-arvoon, ei pörssinoteerattujen yritysten lukumäärässä kuitenkaan ole suurta eroa. Yhdysvaltalaiset pörssiyritykset ovat siten keskimäärin selvästi suurempia kuin mannereurooppalaiset. Tämä markkinarakenteiden ero saattaa johtua siitä, että Euroopan hyödykemarkkinat ovat ainakin 1990-luvulle saakka olleet selvästi vähemmän yhtenäiset kuin Yhdysvalloissa ja markkinarakenteet ovat monilla aloilla noudatelleet kansallisia rajoja.

Vaikka euroalueen rahoitusmarkkinat ovat siis suuret, ne eivät ole niin yhtenäiset eivätkä tehokkaat kuin ne voisivat olla – tai niin yhtenäiset ja tehokkaat kuin niiden pitäisi olla, jotta ne palvelisivat euroalueen talouselämää

Taulukko 2.2. Osakemarkkinoiden kapitalisaatio eri talousalueilla

	Helmikuu 2003, mrd. euroa
USA	9781
Euroalue	3066
Japani	1891
(Lontoo	1509)

Lähteenä on World Federation of Exchanges (FIBV).

Taulukko 2.3. Markkinoilla olevien velkapaperien kanta eri talousalueilla

	Vuoden 2001 lopussa, mrd. euroa
USA	14587
Euroalue	7728
Japani	8095

Lähteenä on Euroopan keskuspankki.

tehokkaimmalla mahdollisella tavalla. Tämä johtuu nykyisin vallitsevan näkemyksen mukaan pitkälti juuri markkinoiden pirstoutuneisuudesta ja korkeista liiketoimikustannuksista. Esimerkiksi arvopaperikeskukset toimivat Euroopassa edelleen pääosin kansallisella pohjalla, ja yksinkertaisten linkkien puuttuessa arvopaperikeskusten väliltä kansainvälinen arvopaperikauppa edellyttää erityisiä custody-järjestelyjä jopa euroalueen sisällä. Itse asiassa erillisiä arvopaperikeskuksia on euroalueella selvästi enemmän kuin valtioita, koska monissa maissa erityyppiset arvopaperit ovat eri järjestelmissä.

Rahoitusmarkkinoiden yhtenäistäminen onkin ymmärrettävästi erittäin korkealla Euroopan unionin prioriteeteissa. EU:n vuonna 1999 vahvistama rahoituspalvelujen toimenpidesuunnitelma (Financial Services Action Plan, FSAP) pyrkii osaltaan yhtenäistämään ja kehittämään arvopaperimarkkinoiden infrastruktuuria varsinkin säätelyn osalta. Toimenpidesuunnitelmaan kuuluu kolme strategista tavoitetta: yhtenäiset rahoituksen tukkumarkkinat, avoimet ja turvalliset rahoituspalvelujen vähittäismarkkinat ja markkinoiden ajanmukainen sääntely ja valvonta. Näiden tavoitteiden rinnalla on kiinnitetty huomiota myös verokysymyksiin, varsinkin säästöjen ja työeläkkeiden verotuksen yhtenäistämiseen. Alkuperäiseen toimenpidesuunnitelmaan ei sisällytetty konkreettisia erityistoimia arvopaperien selvitys- ja toimitusjärjestelmien tehostamiseksi, mutta EU:n komissio on valmistellut jatkotyötä myös tällä lohkolla.

Muutamissa tutkimuksissa on pyritty arvioimaan rahoitusmarkkinoiden yhtenäistymisen mahdollisia vaikutuksia. Erään tuoreen, EU:n komission tilaaman tutkimuksen mukaan reaaliset pääomakustannukset voisivat alentua noin 0.4-0.5 prosenttiyksiköllä, kun osakkeiden ja yritysobligaatioiden markkinat tehostuisivat markkinoiden yhtenäistymisen seurauksena (London Economics 2002). Tällä olisi tutkimuksessa käytettyjen mallien mu-

Taulukko 2.4. Pörssinoteerattujen yritysten lukumäärä eri talousalueilla

	Helmikuu 2003	
	Kaikki	Ulkomaisia
USA	6473	890
Euroaluc	6380	836
Japani	3470	83
(Lontoo	2808	412)

Euroalueen pörssien ulkomaiset listautumiset sisältävät listautumiset muista euromaista. Lähteenä on FIBV.

kaan noin yhden prosentin pysyvä vaikutus kokonaistuotannon tasoon koko EU-alueella. Mikäli edut ovat tätä luokkaa, voi markkinoiden yhdentymisen ja tehostumisen odottaa aikaa myöten vauhdittavan arvopaperistumista ja arvopaperikauppaan perustuvan rahoituksen kasvua myös euroalueella. Miten nopeasti tämä toteutuu, riippuu kuitenkin todennäköisesti paljon paitsi markkinoiden rakenteesta, myös yleisen talouskehityksen suotuisuudesta euroalueella lähivuosina.

KANSAINVÄLISTEN PÄÄOMANLIIKKEIDEN TEHTÄVÄT JA EUROOPAN TARPEET

Taloustieteen peruslähtökohtien mukaan markkinoiden yhtymisestä saatavat hyödyt ovat sitä suuremmat, mitä enemmän eri markkinoilla toimivien osapuolten tarpeet poikkeavat toisistaan. Tästä näkökulmasta voi arvioida, että rahoitusmarkkinoiden onnistuneen globaalin yhdentymisen hyödyt olisivat vielä paljon suuremmat kuin mitä alueellisella integraatiolla voidaan saavuttaa. Kansainväliset pääomanliikkeet ovatkin kasvaneet viime vuosikymmeninä erittäin nopeasti, selvästi nopeammin kuin tavaroiden tai palvelujen maailmankauppa. Kasvu johtuu tietenkin viime kädessä juuri niistä suurista taloudellisista eduista, joita kansainvälinen rahoitus tarjoaa sekä sijoittajille että rahoituksen käyttäjille.

Kansainvälisten rahoitusmarkkinoiden toiminnasta saatavat hyödyt ovat vaikeasti konkretisoitavissa. Yksi syy tähän on se, että nämä hyödyt ovat kovin moniulotteisia. Mielenkiintoista kuitenkin on, että kaikki taloustieteen tärkeimmät tutkimushaarat korostavat kansainvälisten rahoitusmarkkinoiden toiminnasta saatavia hyötyjä, mutta kukin omasta hyvin erilaisesta näkökulmastaan, eikä näillä näkökulmilla näytä olevan kovinkaan paljon tekemistä keskenään. Vaikka makrotaloustieteen, rahoitusteorian ja toimialataloustieteen näkökulmat kansainvälisiin pääomanliikkeisiin ovatkin hyvin erilaisia, kukin näistä tutkimushaaroista korostaa omalla tavallaan sitä, miten läheisesti esimerkiksi Euroopan talouspolitiikan tärkeimmät kysymykset liittyvät kansainvälisten rahoitusmarkkinoiden toimintaan, siis Euroopan rooliin sijoittajana ja sijoituskohteena.

Euroopan talouspolitiikan tärkeimpiä haasteita 2000-luvun ensi vuosikymmenellä ovat muun muassa

- työllisyyden parantaminen,
- väestön ikääntymisestä aiheutuvien kustannusten hallinta,
- kehitysmaiden ja teollisuusmaiden välisen kuilun kaventaminen, ja

- rahoitusmarkkinoiden vakauden ja tehokkuuden lisääminen.

Kaikki nämä haasteet ovat rahoitusmarkkinakysymyksiä, ja pitkälti nimenomaan kansainväliseen rahoitukseen liittyviä. Niinpä esimerkiksi kestävän, työllistävän kasvun luominen Euroopassa riippuu paljon siitä, miten tehokkaasti rahoitusmarkkinat etsivät ja rahoittavat innovaatioita ja kuinka tehokkaasti ne tukevat kasvavien yritysten toimintaa. Väestön ikääntymisen hallinta puolestaan vaatii tuottavia investointeja ja ajoissa tapahtuvaa eläkkeiden rahastointia. Globaalilla tasolla yksi lupaavimmista keinoista kehitysmaiden ja rikkaiden maiden välisen kuilun kaventamiseen on teknologiaa ja muuta osaamista siirtävä pääomasijoitustoiminta kehitysmaihin. Rahoitusmarkkinoiden vakauden ja tehokkuuden lisääminen vaatii kansainvälistä näkökulmaa rahoitusvalvonnan, säätelyn sekä rahoituspalvelujen kaupan kehittämiseen – sekä EU:n tasolla että globaalisti.

PÄÄOMANLIIKKEET JA VAIHTOTASEET

Tarkastellaan Euroopan kansainvälisen rahoitusmarkkinaroolin perusteita ensin makrotaloudellisesta näkökulmasta. Tällöin kansainvälisten rahoitusmarkkinoiden tehtäväksi nähdään mahdollistaa rahoitusylijäämien ja -alijäämien kansainvälinen kauppa. Kansainväliset pääomanliikkeet tekevät toisin sanoen mahdolliseksi sen, että toiset kansantaloudet voivat investoida enemmän kuin säästävät, samaan aikaan kun toiset kansantaloudet puolestaan säästävät enemmän kuin investoivat. Nämä säästöjen ali- ja ylijäämät heijastuvat maiden vaihtotaseissa. Tämän tuottavuutta ja säästämistä korostavan vaihtotaseanalyysin pioneereja oli Blanchard (1983).

On useita syitä, miksi intertemporaalisen vaihdannan mahdollisuus voi hyödyttää kaikkia siihen osallistuvia maita. Esimerkiksi pääoman puutteesta kärsivän kehitysmaan näkökulmasta ulkomaisen pääoman tuonti voi tarjota mahdollisuuden nopeampaan taloudelliseen kasvuun kuin mihin sen oma säästäminen riittäisi. Teorian mukaan pääoman tuonti tällaiseen maahan alentaa siellä investointien tuottovaatimusta ja samalla korottaa työn tuottavuutta ja reaalipalkkoja. Toisaalta pääomantuonti voi olla tehokas ratkaisu myös rikkaalle teollisuusmaalle, jossa avautuu uusia, muita maita paremmin kannattavia investointimahdollisuuksia. Näin on esimerkiksi usein selitetty Yhdysvaltojen nopeaa velkaantumista ulkomaille viime vuosina. Tällöin on ajateltu, että Yhdysvalloissa talouden suuren joustavuuden ansioista pystyttäisiin ottamaan uutta teknologiaa kaupalliseen tuotantoon nopeammin kuin muualla.

Yksi keskeinen makrotaloudellinen selitys vaihtotaseen ylijäämille tai vajeille ovat väestömuutokset. Esimerkiksi nyt monissa rikkaissa maissa näköpiirissä oleva väestön ikääntyminen ja väestönkasvun hidastuminen lisää niissä säästämistä, koska väestö haluaa varautua vanhuuden varalle, mutta vähentää samalla investointikysyntää näissä samoissa maissa, koska työvoiman tarjonta vähenee. Niille maille, joissa ikääntyminen tapahtuu muita maita aikaisemmin tai nopeammin, pääomanvienti tarjoaa mahdollisuuden hankkia ulkomaista varallisuutta, jota voidaan myöhemmin käyttää tuonnin maksamiseen ikääntyneen sukupolven tarpeisiin. Tuloksena on tasaisemmin kasvava kulutusvirta kuin ilman ulkomaisten sijoitusten käyttöä. Tällaisia talouksia, joille demografisista syistä johtuva pääomanvienti voi olla kannattavaa, ovat erityisesti Euroopan ja Japanin nopeasti ikääntyvät kansantaloudet.

YK:n ennusteiden mukaan yli 65-vuotiaiden osuus Euroopan maiden väestöstä nousisi vuoteen 2050 mennessä nykyisestä noin 15-18 prosentista noin 30-34 prosenttiin. Muista maailmantalouden voimakeskuksista Japanissa kehitys on samansuuntainen, mutta tapahtuu vielä aikaisemmin kuin Euroopassa. Japanin työikäisen väestön määrä on itse asiassa jo kääntynyt laskuun. Yli 65-vuotiaiden väestöosuuden oletetaan nousevan siellä yli 30 prosentin jo vuoteen 2030 mennessä ja jatkavan nousua seuraavinakin vuosikymmeninä. Toisaalta Yhdysvalloissa väestön ikääntymisen ei odoteta olevan läheskään yhtä dramaattista kuin Euroopassa tai Japanissa. Siellä yli 65-vuotiaiden osuuden ennustetaan kasvavan nykyisestä noin 13 prosentista noin 20 prosenttiin vuoteen 2030 mennessä ja stabiloituvan sitten. Muualla maailmassa, kuten kehitysmaissa, joissa maailman väestön ylivoimainen enemmistö asuu, odotetaan myös tapahtuvan melkoista ikärakenteen muutosta. Ikääntyneiden osuus kehitysmaiden väestöstä on kuitenkin vielä melko pieni (vain noin 6%) ja se alkaa kasvaa syntyvyyden alenemisesta ja elinajan

Taulukko 2.5. Maailman väestön ikääntyminen YK:n väestöennusteen mukaan

	Ikääntyneiden (yli 65v.) osuus		
	1950	2000	2050
Eurooppa	8%	15%	29%
Yhdysvallat	8%	12%	21%
Japani	5%	17%	36%
Koko maailma	5%	7%	19%

Lähteenä on YK.

pitenemisestä huolimatta vasta paljon myöhemmin kuin Euroopassa tai Japanissa.

Koska maailman pitkän aikavälin väestötrendit ovat näin vahvasti epäsymmetriset, ne luovat periaatteessa suuria mahdollisuuksia kaikkia osapuolia hyödyttäville pääomanliikkeille. Käytännössä rikkaimpien maiden pääomanvienti kehittyviin maihin on kuitenkin ollut vähäisempää kuin taloudellisten fundamenttien perusteella voisi olettaa, ja 2000-luvun alkuvuosina rikkaiden teollisuusmaiden yhteenlaskettu vaihtotase on itse asiassa ollut alijäämäinen, mikä osoittaa, että ne ovat nettomääräisesti laskettuna vetäneet sijoituksiaan pois kehittyvistä maista. "Väärään" suuntaan kulkeva pääomavirta on luonteeltaan osaksi suhdanneluontoinen palautumisilmiö. Toisaalta se osoittaa selvästi, kuinka tärkeää kehittyville maille on luoda toimillaan suotuisat puitteet investoinneille ja ulkomaisille sijoituksille.

Suurin välitön syy rikkaiden maiden yhteenlasketun vaihtotaseen alijäämäisyyteen on tunnetusti Yhdysvaltojen valtava vaihtotasevaje. Euroopassa ja Japanissa väestö ikääntyy huomattavasti nopeammin kuin Yhdysvalloissa, eli tällä perusteella on luontevaa, että vaihtotaseet Euroopassa ja Japanissa ovat ylijäämäisempiä kuin Yhdysvalloissa. Voidaan kuitenkin kysyä, ovatko nykyisen suuruiset erot kehittyneiden maiden vaihtotaseissa kuitenkaan sopusoinnussa taloudellisten perustekijöiden kanssa, varsinkin kun Yhdysvaltain alijäämää lisää sen valtiontalouden suuri alijäämä. Pitäisikö siis sekä Yhdysvaltain että Euroopan säästää enemmän – ja voisiko Japani puolestaan pyrkiä jonkin verran pienempiinkin ylijäämiin?

Yksi mahdollinen selitys euroalueen vaihtotaseen "puuttuvalle" ylijäämälle on, että useimpien euroalueen maiden eläkejärjestelmä on "pay-as-you-go" -tyyppinen eikä rahastoiva. Tällä on ikääntyvässä taloudessa säästämistä vähentävä vaikutus, koska nykyiset sukupolvet voivat perustaa vanhuuden turvansa ainakin osaksi syntymättömien ikäluokkien verottamiseen sen sijaan, että säästäisivät itse. Seurauksena on, että vaihtotase muodostuu

Taulukko 2.6. Säästäminen, investoinnit ja vaihtotase eri alueilla

	Prosenttia BKT:sta vuonna 2001		
	Säästäminen	Investoinnit	Vaihtotase
Euroalue	21%	21%	0%
Yhdysvallat	14%	18%	-4%
Japani	27%	25%	2%

Lähteenä on International Financial Statistics.

väestön ikääntymisvaiheessa heikommaksi kuin rahastoivan eläkejärjestelmän oloissa. Toinen selitys on monien euromaiden julkistalouksien sitkeä alijäämäisyys, mikä tietysti ongelmana kietoutuu yhteen eläkerahastointikysymyksen kanssa.

Se, että euroalueen maat eivät enempää varaudu väestönsä ikääntymiseen säästämällä, haittaa sekä näitä maita itseään että myös koko maailmantaloutta. Euroalueen omasta näkökulmasta haitat johtuvat siitä, että alue ei hyödynnä kansainvälisiä sijoitusmahdollisuuksia kulutuksen kasvun tasaamiseen, vaan tyytyy sen sijaan hitaampaan kulutuksen kasvuun tulevaisuudessa kuin mikä olisi sijoitusten avulla mahdollista saavuttaa. Mitä vähemmän ikääntymiseen varaudutaan rahastoimalla, sitä enemmän verotusta on ilmeisesti tulevaisuudessa kiristettävä, mikä heikentää euroalueen talouden suorituskykyä.

Kehitysmaiden näkökulmasta teollisuusmaiden heikko vaihtotase on ongelma kahdella tapaa. Ensinnäkin teollisuusmaiden pääomanvienti kehitysmaihin jää vähäisemmäksi ja kehitysmaiden talouskasvu näin ollen hitaammaksi kuin mitä se voisi olla. Toiseksi pääomanviennin vähyys vahvistaa teollisuusmaiden reaalista valuuttakurssia, koska globaali kokonaiskysyntä kohdistuu nyt enemmän teollisuusmaihin kuin siinä tapauksessa, että investointeja kehitysmaissa saataisiin laajemman pääomanviennin kautta vauhditettua. Samalla kehittyvien maiden reaalinen valuuttakurssi on heikompi kuin se voisi olla. Pääomanviennin vähyys rikkaista maista kehittyviin maihin on näin osaltaan syynä siihen, että kehittyvien maiden vientituotteiden hintakehitys on jo kauan ollut heikkoa. Kolmannen maailman jo kauan huonona pysynyt vaihtosuhde- ja reaalitulokehitys palautuu siten ainakin osittain teollisuusmaiden pääomanviennin vähyyteen.

Pääomanliikkeillä on aina vastineensa hyödykemarkkinoilla ja tätä kautta myös kytkentä kauppapolitiikkaan. Pääomanviennin kohteena olevan maan on tuotava lainaamansa resurssit viime kädessä aina hyödykkeinä, kauppataseen alijäämän muodossa. Toisaalta pääomaa vievä maa, esimerkiksi rikas, ikääntyvä teollisuusmaa, sijoittaa pääomaa ulkomaille viime kädessä siksi, että voisi kuluttaa tavaroita ja palveluja enemmän kuin itse tuottaa siinä vaiheessa, kun sijoituksia tai niiden tuottoa aikanaan käytetään eläkkeiden rahoitukseen. Tällainen maa ylläpitää siis kauppataseen ylijäämää "nyt" voidakseen ylläpitää alijäämää "myöhemmin".

Kauppapoliittisten näkökohtien vuoksi pääomanliikkeiden ja kauppavirtojen liittyminen toisiinsa ei ole käytännössä yhtä selvää ja ongelmatonta kuin teoriassa. Esimerkiksi monet velkaantuneet kehitysmaat katsovat, että niiden on vaikea lyhentää velkojaan juuri siksi, että niiden vientituotteita syr-

jitään maailmanmarkkinoilla, eivätkä ne siksi kykene aikaansaamaan niitä kauppataseen ylijäämiä, joita velkojen takaisin maksu edellyttää. Kaupan esteet eivät ainoastaan haittaa jo olemassa olevien velkojen hoitamista, vaan myös uusien sijoitusten tekemistä, koska kehitysmaita syrjivä kauppapolitiikka vähentää kehitysmaiden luottokelpoisuutta ja sijoitusten kannattavuutta kehitysmaissa.

PÄÄOMANLIIKKEET JA RISKIEN TASAUS

Toinen keskeinen näkökulma kansainväliseen rahoitukseen on portfolioteoreettinen. Sen merkityksen perusteluksi riittää, kun todetaan, että rahoitusresurssien siirtäminen maasta toiseen, jota makrotaloudellinen näkökulma korostaa, selittää todellisuudessa vain hyvin pienen osan kansainvälisistä sijoituksista. Ristikkäiset sijoitukset eri maiden välillä ovat monin verroin suuremmat kuin ne nettovirrat, joilla vaihtotaseiden alijäämät viime kädessä rahoitetaan ja joiden tuloksena pääomaa sanan varsinaisessa merkityksessä siirtyy maasta toiseen. Tämä käy selvästi ilmi esimerkiksi oheisista, maiden välisiä sijoituspositioita kuvaavista tiedoista.

Rahoitustutkimus omasta näkökulmastaan korostaa ristikkäisten sijoitusten hyödyllisyyttä, koska ne tarjoavat mahdollisuuden sijoitusriskien ta-

Taulukko 2.7. Talousalueet sijoittajina

	Vuoden 2001 lopussa, mrd. euroa		
	Euroalue	USA	Japani
Suorat sijoitukset (ulos)	1860	1549	349
Portfoliosijoitukset	2500	2365	1499
Kaikki sijoitukset	7459	6946	3349

Lähteenä on Euroopan komissio.

Taulukko 2.8. Talousalueet sijoituskohteina

	Vuoden 2001 lopussa, mrd. euroa		
	Euroalue	USA	Japani
Suorat sijoitukset (sisään)	1329	1481	58
Portfoliosijoitukset	3183	4785	774
Kaikki sijoitukset	7612	9130	1768

Lähteenä on Euroopan komissio.

saamiseen eri maiden kesken. Tähän kiinnittivät huomiota Levy ja Sarnat klassisessa tutkimuksessaan (1970) jo kolme vuosikymmentä sitten. Tyypillisten laskelmien mukaan esimerkiksi yhdysvaltalainen sijoittaja pystyisi jopa puolittamaan osakesijoituksiinsa liittyvän riskin hajauttamalla portfolionsa kansainvälisesti sen sijaan että tyytyisi sijoittamaan vain yhdysvaltalaisten yhtiöiden osakkeisiin. Nämä hajauttamisesta saatavat hyödyt saavutetaan tinkimättä portfolioiden tuottovaatimuksista.

Vaikka kansainvälisten pääomanliikkeiden bruttovirrat todella ovatkin paljon suuremmat kuin vaihtotaseita rahoittavat nettovirrat, tiedetään myös, että sijoitusten kansainvälinen hajauttaminen on käytännössä paljon vähäisempää kuin portfolioteorian perusteella voisi odottaa. Esimerkiksi 1990-luvun puolivälissä yhdysvaltalaisten sijoittajien osakesijoituksista vain noin 10 prosenttia oli sijoitettuna ulkomaille. Kaikki käytettävissä olevat tiedot viittaavat siihen, että myös euroalueella toimivien sijoittajien portfolioissa on miltei samaa suuruusluokkaa oleva "kotimaaharha". Tätä pidetään toisena kansainvälisten rahoitusmarkkinoiden suurista paradokseista, melkein yhtä tärkeänä kuin nettopääomanliikkeiden pienuutta, josta edellä oli puhe.

Tutkimusten mukaan kansainvälisen hajautuksen vähyys johtuu ainakin seuraavista syistä. Ensinnäkin transaktiokustannukset haittaavat varojen tehokasta hajautusta varsinkin piensijoittajien tapauksessa. Toiseksi ulkomaisiin sijoituksiin liittyy suurempia informaatio-ongelmia kuin kotimaisiin sijoituksiin. Portfolioteorian näkökulmasta nämä ongelmat estäisivät sijoitusriskien tehokkaan tasaamisen maailman eri osissa toimivien sijoittajien kesken ja siten osaltaan ovat syynä arvopaperisijoitusten suuriin riskeihin. Toisaalta on hyvin todennäköistä, että hyödykemarkkinoiden lisääntynyt yhdentyminen viime vuosina on osaltaan synkronisoinut sijoitusriskejä eri markkinoilla ja tätä kautta vähentänyt sijoitusten kansainvälisestä hajauttamisesta saatavissa olevaa hyötyä.

Euroalueen kannalta sijoitusten globaali hajauttaminen on itse asiassa sijoitusten kansainvälistymisen toinen vaihe: ensimmäinen on hajauttaminen euroalueen sisällä, johon on rahaliiton ansiosta aikaisempaa huomattavasti paremmat mahdollisuudet nyt, kun valuuttakurssiriskit ovat poistuneet. Näin esimerkiksi euroalueen institutionaaliset sijoittajat (kuten eläkerahastot), jotka eivät ota valuuttakurssiriskiä, ovat voineet hajauttaa sijoituksiaan yhteisvaluutta-alueen sisällä ennen näkemättömällä tavalla. Tällä on Suomen pääomanliikkeiden kannalta ollut viime vuosina aivan ratkaiseva merkitys. Euroalueen arvopaperimarkkinoiden pirstoutuneisuus on kuitenkin tehnyt tämän yhä liian kalliiksi varsinkin piensijoittajien kannalta.

Jos rahoitukseen liittyvät informaatio-ongelmat olisivat vähäiset ja jos rahoitusmarkkinoiden infrastruktuuri olisi hyvin kehittynyt, arvopaperimarkkinat olisivat tehokas muoto rahoituksen välitykseen ja sijoitusriskien hajauttamiseen. Jos nämä ehdot täysin toteutuisivat, tuskin voitaisiin löytää parempaa tapaa yhdistää riskien hajautus, sijoitusten likvidiys, pääoman hinnan kilpailullinen määräytyminen ja monet muut hyvin toimivien markkinoiden tunnusmerkit kuin portfoliosijoituksiin ja arvopaperimarkkinoihin perustuva rahoitus.

Käytännössä sijoitusten "kotimaaharha" osoittaa omalta osaltaan sen, että kansainvälisten arvopaperimarkkinoiden toiminnassa on vielä paljon toivomisen varaa. Ongelmista kertoo myös kehittyvien maiden kansainvälisten joukkolainojen markkinoiden toistuva epävakaus ja aika ajoin tapahtuva kriisiytyminen – viime vuosina esimerkiksi latinalaisen Amerikan ja Venäjän tapauksissa. Kansainvälinen rahoitusviranomaisten yhteistyö onkin viime vuosina tähdännyt etenkin informaatio-ongelmien vähentämiseen ja juridisen infrastruktuurin vahvistamiseen. Esimerkiksi yritysten tilinpäätösinformaation laatu ja vertailtavuus, samoin kuin eri maiden taloustilastojen luotettavuus ja nopeus ovat kansainvälisen yhteistyön painopisteitä. Pelisääntöjen laatiminen valtioiden velkaongelmien varalta on eräs mielenkiintoinen hanke, joka voisi vähentää valtiolainoihin liittyviä informaatio-ongelmia. Euroalueen kannalta globaalien arvopaperimarkkinoiden kehitys on vakaus- ja riskikysymys; kehitysmaiden ja uusien teollisuusmaiden kannalta kysymys on ennen muuta sijoituspääoman hinnasta ja siihen liittyvien riskipreemioiden vähentämisestä.

PÄÄOMANLIIKKEET JA OSAAMISEN SIIRTO

Siinä missä makrotaloudellinen näkökulma kansainvälisiin sijoituksiin korostaa rahoitusylijäämien kauppaa ja portfolionäkökulma puolestaan sijoitusriskien kauppaa, kolmas keskeinen lähestymistapa eli toimialataloustieteen tai liiketaloustieteen näkökulma korostaa teknologian siirron ja yritysosaamisen hyödyntämisen merkitystä kansainvälisissä sijoituksissa.

Tämä näkökulma korostaa sitä, että omistuksessa ei ole kysymys vain varallisuuden hallussapidosta ja sen tuotosta nauttimisesta (joita portfolioteoria korostaa) vaan omistus on aktiivista toimintaa, johon liittyy yritysten toiminnan kontrollia ja kehittämistä. Kansainvälisessä sijoitustoiminnassa korostuu tästä näkökulmasta tietämyksen, kuten johtamistaidon ja tuotantoteknologian, siirtyminen maasta toiseen. Pääomanliikkeet merkitsevät tällöin yritysten kansainvälistymistä ja monikansallisten konsernien syntymistä.

Motiivina tähän on yrityksissä olevan osaamispääoman mahdollisimman tehokas hyödyntäminen. Tätä tekijää korosti jo Servan-Schreiber (1968) tarkastellessaan amerikkalaisten monikansallisten yritysten laajenemista Eurooppaan 1960-luvulla. Nykyaikainen käsitys yritystoiminnasta korostaa niin sanottua hiljaista tietoa, joka vaikuttaa organisaation toiminnassa, mutta jota ei voi helposti ostaa eikä myydä markkinoilla (hiljaisesta tiedosta ks. Polanyi, 1962). Tällöin yrityksen laajentaminen on ainoa tapa saada tämä pääoma sekä kansantalouden että yrityksen itsensä kannalta tehokkaaseen käyttöön.

Kansainväliset suorat sijoitukset ovat 1980-luvulta alkaen olleet erittäin nopeassa kasvusta, ajoittaisista ja lyhytaikaisista takaiskuista huolimatta. Sekä euroalueen että Yhdysvaltojen ulkomaisista sijoituksista noin neljännes on nykyisin suoria sijoituksia. Kansainvälisten suorien sijoitusten luonteeseen kuuluu, että ne ovat käytännössä suurimmaksi osaksi niin sanottua ristikkäiskauppaa, jossa kehittyneimmät maat investoivat toisiinsa. Esimerkiksi EU-alueen suorista ulkomaisista sijoituksista noin puolet on suuntautunut Yhdysvaltoihin ja Japaniin; vastaavasti Yhdysvaltojen suorista ulkomaisista sijoituksista noin puolet on suuntautunut EU-alueelle ja Japaniin. EU-maat ovat sitä paitsi sijoittaneet runsaasti lähialueelleen EFTA-maihin (kuten Sveitsiin) ja USA vastaavasti Kanadaan.

Mielenkiintoista Euroopan suorissa ulkomaisissa sijoituksissa viime vuosina on ollut se, että erittäin suuri osa niistä on suuntautunut rahoitussektoriin. Eurooppalaiset pankit ja finanssikonsernit ovat toisin sanoen perustaneet tytärpankkeja ja muita rahoitusalan yrityksiä sijoitusten kohdemaihin. Myös sijoitukset kehittyviin talouksiin ovat usein tapahtuneet rahoitussektoriin. Esimerkiksi vuosina 1996-2000 tehdyistä EU-maiden suorista ulkomaisista sijoituksista noin neljännes on tehty rahoitustoimialalle, mikä merkitsee tämän toimialan huomattavaa yliedustusta verrattuna rahoitusalan osuuteen esimerkiksi EU:n tai kohdemaiden kokonaistuotannosta.

Suoria sijoituksia pidetään usein erityisen otollisena tapana viedä kehittyviin maihin niiden tarvitsemaa ulkomaista pääomaa. Tähän on useita syitä. Ensinnäkin suorat sijoitukset voivat siirtää yrityskohtaista osaamista, kuten johtamis- ja markkinointiosaamista jne. maihin, joissa sitä ennestään on niukasti. Itse asiassa "hiljaisen tiedon" luonteeseen kuuluu, että suorat sijoitukset saattavat olla lähes ainoa tapa, jolla yritysten sisäinen osaaminen on itse asiassa siirrettävissä. Toiseksi suorat sijoitukset ovat sijoittajan erityisen läheisessä valvonnassa, joten monet niistä informaatio-ongelmista, jotka haittaavat "arm's length" -tyyppistä arvopaperisijoittamista kehittyviin maihin, voidaan suorilla sijoituksilla välttää. Kolmanneksi suorat sijoitukset ovat pääsääntöisesti oman pääoman ehtoisia. Tämä merkitsee sitä, että ne eivät

aiheuta velkataakkaa, ja niistä sijoittajalle kotiutettava osinkovirta periaatteessa vaihtelee suhteessa sijoituksen menestykseen. Velkaongelmien tapaiset riskit ovat siten suorien sijoitusten muodossa tapahtuvan pääomantuonnin tapauksessa kohdemaalle pienemmät kuin velkapääomaa tuotaessa.

Euroopan kannalta suorien sijoitusten kehittäminen on useastakin syystä suotavaa. Jos euroalueen on syytä pyrkiä lisäämään pääomanvientiään ulkomaille ja erityisesti kehittyviin maihin, kuten edellä esitin, suorat sijoitukset tarjoavat tähän keinon, joka voi suureksi osaksi välttää portfoliosijoituksiin (ja interbank-rahoitukseen) liittyvät vakausongelmat. Lisäksi suorat sijoitukset ulkomaille maksimoivat eurooppalaisissa yrityksissä olevasta osaamispääomasta saatavan kaupallisen hyödyn. Kolmanneksi, ne toteuttavat ehkä tehokkaammin kuin mikään muu pääomanviennin muoto osaamisen siirtoa kehittyneemmistä maista kehittyviin maihin ja auttavat kaventamaan maiden välisiä kehityskuiluja. Ulkomaisia suoria sijoituksia tarkastellaan usein aivan liian yksipuolisesti "työpaikkojen viennin" tai "sosiaalisen polkumyynnin" tapaisten otsikoiden alla ottamatta huomioon niiden ainutlaatuista luonnetta verrattuna muihin pääomanliikkeiden lajeihin.

EU:n rahoitusmarkkina- ja kauppapolitiikassa kansainväliset suorat sijoitukset ovat nykyisin etusijalla. Tämä linja on tullut selvästi esiin varsinkin sen jälkeen kun OECD:n piirissä käydyt multilateraalista investointisopimusta (MAI) koskevat neuvottelut päättyivät lokakuussa 1998 käytännössä tuloksettomina. Välittömänä syynä oli, että Ranska vetäytyi näistä pääomanliikkeiden yleistä vapauttamista koskeneista neuvotteluista pian sen jälkeen kun Venäjän kriisi oli ravistellut kansainvälisiä valuutta- ja rahoitusmarkkinoita. Multilateraalisia, investointipolitiikkaa koskevia neuvotteluja on sen jälkeen käyty WTO:n piirissä, mistä sovittiin WTO:n ministerikokouksessa Dohassa marraskuussa 2001. Keskeisiä asioita Dohan kierroksella ovat palvelujen kaupan ohella nimenomaan kansainväliset sijoitukset. Tässä vaiheessa neuvottelut etenevät nopeammin palvelujen kaupan alueella, mutta investointineuvottelut ovat nekin alkaneet. EU:n positio sekä palvelukaupan vapauttamisen että investointien vapauttamisen alueilla on maltillinen ja korostaa maiden suvereenisuutta monissa herkiksi koetuissa kysymyksissä.

EU-mailla on ollut lokakuusta 1999 alkaen yhteinen positio monenkeskisiä WTO:n puitteissa käytäviä investointineuvotteluja varten. Tavoitteena on monenkeskinen kansainvälisten sijoitusten sääntelykehikko, jolla pyritäisiin luomaan vakaa ja ennustettava toimintaympäristö ulkomaisille sijoituksille koko maailmassa. EU:n kannan mukaan sääntelykehikon olisi keskityttävä ulkomaisiin suoriin sijoituksiin, ja lyhytaikaisen pääoman liikkeet olisi tässä vaiheessa jätettävä säännöstön ulkopuolelle. Neuvotteluissa olisi kä-

siteltävä investointimahdollisuuksien avoimuutta ja syrjimättömyyttä, investointisuojaa ja vakaata ja läpinäkyvää yritysilmastoa. Kehikon olisi varmistettava, että kansainväliset sijoitukset voivat tukea kestävää kehitystä, ja säilytettävä sijoitusten kohdemailla mahdollisuus säännellä ulkomaalaisten sijoittajien toimia alueellaan WTO:n perussääntöjen mukaisesti. Tulevan kehikon olisi EU:n position mukaan otettava huomioon kansalaisyhteiskunnan ilmaisemat huolet koskien mm. sijoittajien vastuuta. Lisäksi EU haluaisi, että erimielisyydet ratkaistaisiin valtioiden välillä, jolloin sijoittajat eivät voisi suoraan riitauttaa minkään jäsenmaan lakeja tai sääntelypäätöksiä.

EU:n yhteinen positio WTO-neuvotteluissa heijastaa maltillisuudessaan tiettyä poliittista ja taloudellista realismia. Samalla on toisaalta selvää sekin, että portfoliosijoituksia ja pankkien välisiä rahamarkkinoita koskevia kysymyksiä ei voida kansainvälisessä rahoitusmarkkinapolitiikassa mitenkään sivuuttaa, vaikka WTO-neuvottelut eivät niitä koskekaan. Ne ovat siirtyneet kauppapoliittisista neuvottelupöydistä toisille foorumeille, joissa paino on vakauden edistämisessä. Käytännössä näitä kysymyksiä käsitellään aktiivisimmin tällä hetkellä Kansainvälisen Valuuttarahaston ja BIS:n piirissä, missä EU:n tai edes euroalueen keskinäinen koordinaatio ei ole yhtä tiivistä kuin WTO:ssa. Näkökulma IMF:n ja BIS:n piirissä tehtävässä työssä painottuu kansainvälisten rahoitusmarkkinoiden vakauden valvontaan ja vakautta edistävän sääntelyn valmisteluun ja koordinointiin, ja siksi on luontevaa että siellä keskuspankeilla on keskeinen rooli. Euroopan osalta sekä EKP että kansalliset keskuspankit ovat tärkeitä osapuolia, koska rahoitusmarkkinoiden vakaudesta huolehtiminen on EU:ssakin yhä lähtökohtaisesti kansallisten viranomaisten vastuulla.

ERÄITÄ TULEVAISUUDEN KYSYMYKSIÄ

Samalla kun talousteoria todistaa kansainvälisten pääomanliikkeiden hyödyllisyydestä sekä maailmantaloudelle kokonaisuutena että myös yksittäisten sijoittajien ja rahoituksen käyttäjien kannalta, historiallinen kokemus osoittaa, kuinka suuria ongelmia kansainväliseen rahoitukseen myös liittyy. Ongelmia syntyy varsinkin silloin, kun rahoituksen osapuolina ovat maat, joiden taloudellinen ja yhteiskunnallinen kehitystaso on hyvin erilainen. Syynä siihen, että rahoitusmarkkinoiden kansainvälinen yhdentyminen on ollut vielä kaupan vapauttamistakin vaikeampi prosessi, on että rahoitus sitoo siihen osallistuvat maat vielä läheisempään keskinäiseen riippuvuuteen kuin tavaroiden tai palvelujen kauppa. Siksi rahoitusmarkkinoiden tehokas toiminta vaatii paljon enemmän informaatiota ja vahvemman institutionaalisen

perustan kuin tavaroiden tai palvelujen kauppa. Vain hiukan yksinkertaistaen voi sanoa, että tavarakaupan välttämättömänä edellytyksenä on pelkästään se, että osapuolet kohtaavat toisensa. Rahoitustoimien onnistuminen vaatii tämän lisäksi, että osapuolet myös tuntevat toisensa ja toimivat samojen sääntöjen mukaan.

Kansainvälisten rahoitusmarkkinoiden vakauden varmistaminen olisi vähemmän visainen ongelma, elleivät kansainvälisistä pääomanliikkeistä saatavat potentiaaliset hyödyt olisi niin suuret. Jos näin ei olisi, tarvitsisi vain pohtia, miten pääomanliikkeitä voitaisiin rajoittaa. Tästä kärsisivät erittäin todennäköisesti pitkän päälle kaikki osapuolet – sekä rahoitusylijäämäiset maat että ne, jotka tarvitsisivat ulkomaisia sijoituksia. Kun kansainvälisestä rahoitustoiminnasta saatavat edut kuitenkin ovat ilmeiset, pääomanliikkeet pyrkivät markkinavoimien vaikutuksesta yhä kasvamaan ja myös talouspolitiikan on järkevää edistää markkinoiden toimintaa ja tehokkuutta. Tällöin markkinoiden vakauden varmistaminen nousee keskeiseen asemaan.

Olen edellä pyrkinyt perustelemaan näkemystä, jonka mukaan euroalue ei ole vielä läheskään käyttänyt loppuun niitä mahdollisuuksia, joita kansainvälisen rahoitusmarkkinaroolin kehittäminen sen taloudelle tarjoaisi. Euroalueen sisäisen rahoitusmarkkinaintegraation loppuunsaattamisen lisäksi keskeinen edellytys edistykselle on rahoitusmarkkinoiden globaalin toimintaympäristön kehittäminen vakaaksi, läpinäkyväksi ja kansainvälisesti tasapuoliseksi. Käytännössä tämä tarkoittaa vakaata raha- ja talouspolitiikkaa, transparenssia painottavaa rahoitusmarkkinapolitiikkaa ja kansainvälistä sääntely-yhteistyötä "tasaisen pelikentän" rakentamiseksi kansainvälisille sijoittajille ja rahoituksen tarvitsijoille.

Euroalue ei globaaleilla rahoitusmarkkinoilla voi toimia yksin edes vakauden edistämiseksi, vaan monenkeskinen yhteistyö maailman suurten rahoitusmahtien ja pienempienkin rahoituskeskusten kanssa on välttämätöntä. Jotta euroalue voisi tehokkaasti omalta osaltaan osallistua kansainvälisten markkinoiden toiminnan kehittämiseen, sen on kuitenkin varmistettava sisäinen toimintakykynsä ja riittävä koordinaatio kansainvälisissä esiintymisissä. Tämä merkitsee ainakin Euroopan oman rahoitusvalvonnan tiivistämistä nykyisen melkoisen hajanaisuuden vähentämiseksi ja myös tiiviimpää eurooppalaista yhteistyötä niillä globaaleilla foorumeilla, joilla markkinoiden kehitystyötä nykyisin tehdään.

KIRJALLISUUTTA

Blanchard, O. (1983): Debt and the Current Account in Brazil, teoksessa P. Armella (toim.): *Financial Policies and the World Capital Market: The problem of Latin American Countries*, University of Chicago Press.

European Union Foreign Direct Investment Yearbook 2001, European Communities 2002.

Levy, H. ja Sarnat, M. (1970): Diversification, Portfolio Analysis and the Uneasy Case for Conglomerate Mergers, *Journal of Finance*, vol. 25.

London Economics (2002): *Quantification of the Macro-Economic Impact of Integration of EU Financial Markets*, Report to the European Commission – Directorate-General for the Internal Market, Lontoo.

Polanyi, M. (1962): *Personal Knowledge: Towards a Post-Critical Philosophy*, University of Chicago Press.

Servan-Schreiber, J. (1968): *The American Challenge*, Athenaeum, New York.

3. Economic Integration and Openness in Europe and East Asia

Ray Barrell and Amanda Choy*

OVERVIEW OF THE PAPER

The paper discusses European and East Asian Integration, and makes some comparisons between developments in the two regions. We stress that the process of integration is a political process as well as an economic one, and that the stimulation of competition is more important than the generation of trade in achieving the economic and political objectives that the Europeans have set themselves. It is therefore important before comparing integration in Europe and East Asia to note that the political agenda in the two regions may not be the same.

European integration progressed from a custom union in the 1950s to a single market in the 1980s and finally to a currency union in 1999 and 2002. Over the course of integration, the EU benefited particularly from trade expansion and economies of scale from increasing internal specialisation. In addition it is clear that productivity growth increased as a result of rising competition as economic integration increased. We look at the evolving patterns of trade and competitiveness, and discuss the literature on the role of trade and competition in European economic integration. We also examine the impact of integration on foreign direct investments (FDI) and note that the existence of trade flows may reflect barriers to FDI. The increasing role of FDI by European firms within Europe has helped drive increased competition, but it reflects the movements of firms to markets, rather than the movements of firms to low cost locations, which is more common in East Asia. The impacts of these two forms of relocation of activity are noticeably different.

* Ray Barrell is Senior Research Fellow and Amanda Choy Research Officer at the National Institute of Economic and Social Research (NIESR). This paper was presented at the ASEM Symposium on Multilateral and Regional Economic Relations, Tokyo, March 25th 2003. The authors would like to thank the organizers, their hosts the Japanese Ministry of Foreign Affairs, Governor Jens Thomsen, Professor Pang Eng Fong and the other participants in the session for their valuable comments. The errors remain authors.

We also look at integration in East Asia, and comment on the East Asian Crisis of the late 1990s. We discuss the evolving pattern of trade in East Asia and its determinants, surveying the recent literature on gravity models and trade. This region is much more externally orientated than is Europe, suggesting that any gain from trade following an East Asian integration would be substantially smaller. Furthermore, many of the trade barriers within East Asia, especially in terms of tariffs, have already been removed. This has happened through multilateral platforms such as World Trade Organisation (WTO), Asia Pacific Economic Cooperation (APEC) and Association of South East Asian Nations (ASEAN) as well as through free trade agreements (FTAs) between individual countries. Hence the gains of scale that can come from specialisation have probably also been reaped.

There has been renewed discussion in East Asia of the case for forming a currency union amongst the countries of the region. We look at the potential for a currency union in East Asia in the light of lessons to be learnt from the successful process of monetary integration in Europe. Patterns of competition matter when designing regional integration schemes and assessing exchange rate arrangements. The European economies have been increasingly predominantly competing with each other as competition increases through higher degree of integration. Only a few Asian economies, mostly the Newly Industrialised Economies (NIEs), compete directly with each other with competition already very intense on the export platform. Most other Asian economies compete with countries outside Asia – e.g. ASEAN with Latin Americas, Japan with Europe – suggesting that further integration within Asia might not raise the degree of competition as much as in Europe. We look at measures of competitiveness, and stress the importance of evaluating the group of countries and firms with whom competition takes place.

THE PROCESS OF EUROPEAN INTEGRATION

The foundation of the EU stems from the creation of a Custom Union in 1957 with an aim to enable free movement of factors of production within the region. The initial members were Germany, France, Italy, the Netherlands and Belgium, and the group emerged out of cooperation agreements on coal and steel. Common institutions in a number of policy areas were created while explicit objectives for further integration in other complementary policy areas were agreed. The major initial innovation was the reduction of internal tariff barriers with the objective of removing them entirely, and the creation of a common external tariff. As the members were either low external tariff coun-

tries (Germany, the Netherlands) of high external tariff countries (France, Italy) the obvious compromise was to take the average of these countries. However, as the union evolved it was increasingly clear that non-tariff barriers to trade within the union reduced the gains that were available from greater integration, and increasingly the union moved toward the creation of a single market.

Two sets of non-tariff barriers were identified, and the process of removing them was slow. Barriers can be the result of the existence of different currencies, flexible exchange rates and different financial structures, and hence the process of monetary integration was seen as important. Other non-tariff barriers are associated with the restrictions on the movements of factors of production and the reduction of competition that results from different national regulations and legislation on product standards and the control of production. These can and have been used as effective barriers by countries within the union in order to protect their own industries as well as their consumers, and the Single Market Programme (SMP) which was agreed on in the mid 1980s and implemented from 1986 onwards was designed to remove these barriers. The evidence on the importance of openness and integration for the growth process is surveyed at length in Pain (2002). It is relatively clear from a wide range of studies that freer trade and smaller non-tariff barriers can enhance growth, but that political as well as economic conditions have to be supportive of the growth process[1].

The union increased in scale over the 1970s and 1980s, with the UK, Ireland and Denmark joining in 1973, Greece in 1981 and Portugal and Spain in 1986. This presented the union with additional problems, as the distribution of incomes within the union in the late 1980s was considerably greater than in 1970. In addition, the initial members all shared a legal tradition based on statute law, whilst the UK and Ireland had strong common law traditions, and these differences introduced new types of barriers to competition that would have to be dealt with. The accession of Austria, Sweden and Finland in 1995 presented fewer problems, although again legal traditions and industrial structures were diverse, and these presented barriers to greater competition. The evidence for the UK suggests that the gains from accession were mainly from increased competition and economies of scale. The gains from increased specialisation and comparative advantage were much less. Gaisorek et al (2002) calculate that membership of the EU probably increased UK GDP between 1973 and 1985 by up to 3 percent. The gains from increased specialisation and comparative advantage contributed no more than 1 percent of this. This scale of gains would be available for any new

member, and they could be greater for recent entrants such as Finland who will also reap the dynamic benefits of the Single Market Programme.

Various arrangements for closer monetary co-operation including short term swap arrangements, long term balance of payments support and the European Monetary Co-operation Fund was formed in the 1970s. Co-operation was significantly strengthened by the 1980s through the establishment of an increasingly sophisticated European Monetary System (EMS). The Exchange Rate Mechanism (ERM) of the EMS, which started in 1979, involved a commitment to hold exchange rates within bands, and realignment of the central rate was managed by agreement. During the 1980s realignments became less common, but the system came under strain in the early 1990s in part because of difficulties associated with late UK membership, and also because the impact of German unification. Germany was at the core of the system, and set interest rates in relation to German targets, and unification increased demand there, requiring a rise in rates that was not needed in other European countries. The ERM was an important part of the process of removing barriers in Europe, and its transformation into full monetary union is potentially important.

Real convergence of the European economies was strengthened with the creation of a single market in the late 1980s and early 1990s. Through the Single Market Programme, trade barriers including non-tariffs barriers began to be eliminated while a common regulatory framework was created. The Single Market was designed to allow capital, goods, services and people to move freely and allowed firms as well as investors to develop Europe wide strategies. Allen, Gasiorek and Smith (1998) found that the SMP has increased the intensity of competition within Europe and has led to industrial restructuring as price-cost margins were reduced. The SMP was also found to be strongly welfare increasing for all EU economies. Barrell and Pain (1997b) suggest that many of the gains from the SMP came from the transfer of technology associated with FDI and the relocation of firms within Europe that the SMP stimulated.

TRADE PATTERNS IN EUROPE

The early part of the process of European regional integration was based on the removal of barriers to trade, and it is not surprising that the intensity of trade within the region rose between 1960 and 1980. However, the second phase of integration involved the removal of barriers to competition, and the impact of the SMP on patterns of trade is less clear. Increased competition can

lead to greater specialisation and hence more trade, or increased price competition by making markets actually or potentially more contestable. Greater price competition may not impact directly on trade and location but may involve the reduction of profit margins and the greater efficiency of use of factors of production. In addition the removal of barriers to capital mobility could potentially reduce trade if it allowed firms to locate in more than one country to produce their products, rather than being constrained to stay in their home country and export from there.

Historically, the share of intra-regional trade to total trade in Europe has been high with intra-regional export share among the 15 member countries of the EU standing at 47 per cent even as early as 1960 even though only 6 of them belonged to a customs union. This strong internal orientation increased further with the establishment of the European Free Trade Association (EFTA) in 1960 and the expansion of the European Economic Community (EEC) in 1973. By 1980, the share of intra-regional export to total exports from Europe reached 61 per cent. Despite further expansion of the ECC over the 1980s and 1990s, the level of internal orientation has stabilised as the increase in intra-regional trade from newer members were offset by member countries adopting more open world trade policy. Figure 3.1 plots trade shares with the US and within Europe for the current members of the EU[2].

One reason for the share of intra-regional share to stabilise between 1980 and 2000 was perhaps that many of the gains from trade had been reaped by 1980 and the gains from competition have begun to work through capital mobility. The SMP has had an impact on patterns of trade, but it has not been large. Swann *et al.* (1996) for instance suggest that the implementation of common standards may change import demand, especially within the EU. Commission projects investigating the effects of the SMP have focused on intra-European trade and their results suggest that there has been a mod-

Table 3.1. European Exports

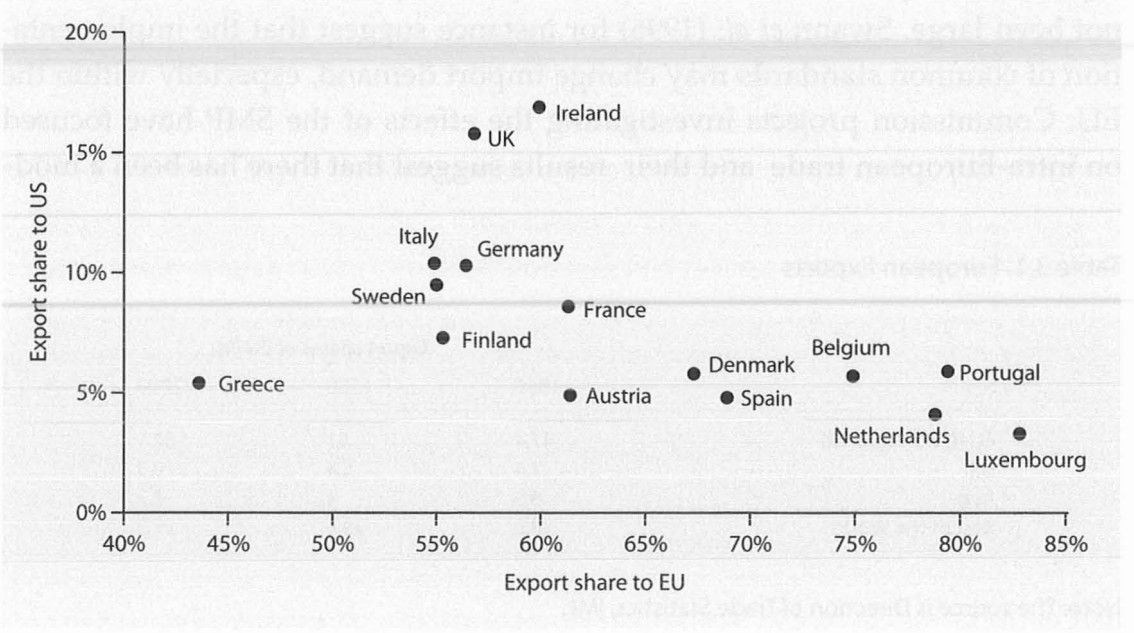
	Export shares of EU (%)		
	1960	1980	2000
EU (Intra-regional)	47.4	61	62.1
US	7.5	5.4	9.3
Asia	6.5	4	7.3
Rest of the World	38.6	29.6	21.3

Note: The source is Direction of Trade Statistics, IMF.

erate increase in intra-European trade created by the SMP (EC, 1997). Barrell and te Velde (2002) found that once they included FDI and technology indicators in a study of import demand in Europe there was little evidence that the creation of the Single Market has directly increased aggregated imports in European countries rather than increasing imports through its effects on location and FDI. They suggest that there are other factors at work, such as increasing specialisation based on R&D that are also having an additional impact on the propensity to import, and these increases should not be mistakenly ascribed to the Single Market Programme.

It can be observed from Figure 3.2 that Finland, Sweden and Ireland stood out from the rest of the EU members in that these 3 countries experienced sustained fall in intra-regional export share between 1960 and 2000 as their trade sector become increasingly open to the world market. Ireland's dependence on trade with the UK declined markedly after accession to the EU in 1973 and especially after the UK decided not to join it in the ERM in 1978. In addition the FDI based growth strategy so successfully pursued by Ireland increased trade with the US in particular as it became an increasingly important platform for US producers in Europe, stimulating Irish exports to the US as well as elsewhere. The Swedish and Finnish experience had much more to do with the changing nature of their own production activities and the rise of the production of telecommunications equipment and mobile pho-

Figure 3.1. European Export Shares in 2000

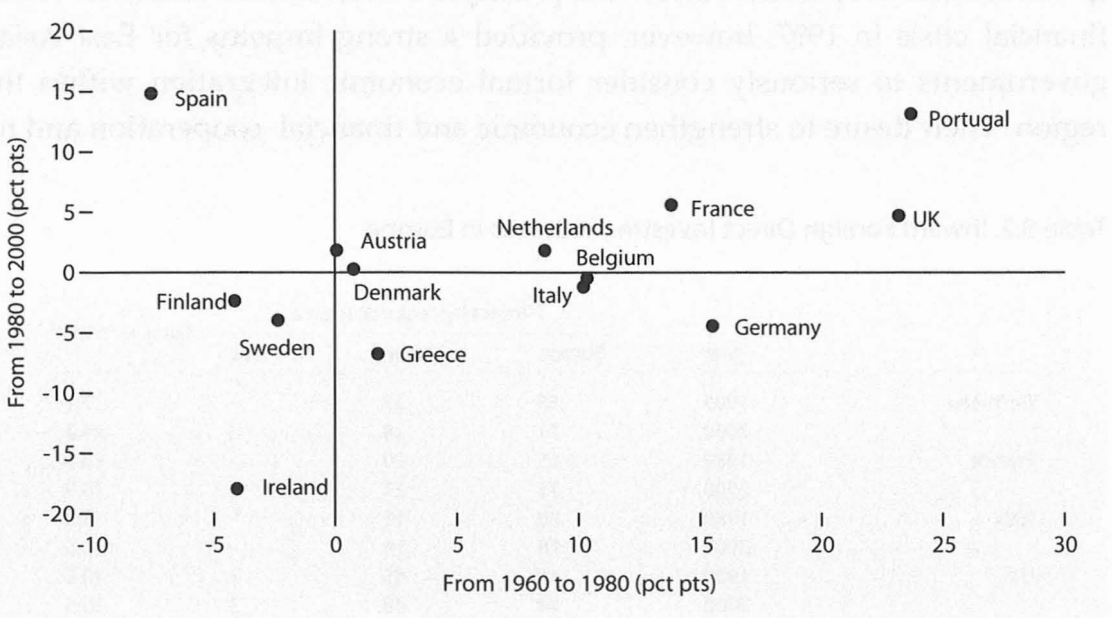


Note: The source is Direction of Trade Statistics, IMF.

nes. Both were late joiners, and had benefited from their decision to deregulate telecoms markets before they joined, giving them first mover advantages in mobile phone technology.

Barrell and Pain (1997b) argue that European integration has had an important effect on the pattern and level of FDI within Europe and this has been a major vehicle for the impact of competition on productivity. Policies pursued collectively by all European governments and the European Commission have helped to stimulate cross-border investments by firms particularly from inside and also from outside the region. The implementation of the Single Market Programme (SMP) has clearly stimulated the mobility of capital and generated a significant change in investment patterns and levels of foreign involvement both in manufacturing and in services. In general there has been an increase in FDI flows within Europe, as can be seen from Table 3.2. The stock of FDI as a percent of GDP rose markedly in the four large European economies between 1989 and 2000, but the increases were particularly marked in France and Germany. The latest data suggest that over a quarter of all manufacturing output in Germany is now produced in foreign owned enterprises, for instance. The increase in stocks of FDI has been particularly concentrated amongst European countries, and this is particularly instructive.

Figure 3.2. Changes in Intra-regional Export Shares within EU



Note: The source is Direction of Trade Statistics, IMF.

Trade and investment take place in part because of the existence of firm specific assets, pieces of knowledge, patents or products that others do not have. Markets for their products can be accessed by trade or by the establishment of additional production facilities in the market that they wish to access. If barriers to trade are reduced but barriers to capital mobility remain, then trade will be enhanced. If barriers to capital mobility are removed then firms may set up plants elsewhere to reap the profits from their specific assets. The choice of strategy once both barriers to trade and factor mobility have been removed depends in part on the economies of scale available within the firm, and as Barrell and Pain (1997a) stress, more modern technologies based on knowledge and computer based organisation may mean that economies of scale have moved from the plant to the firm, and hence large firms may have numbers of plants in order to benefit from scale economies.

EUROPEAN AND EAST ASIAN GROWTH COMPARED

While inter-linkages and inter-dependency between economies within East Asia has become increasingly strong over the years, this informal economic integration was primarily led by multinational corporations (MNCs) which seek to build up marketing and supply chain network to exploit economies of scale from division of labour and specialisation. Until the late 1990s, economic integration on the Asian political platform has been restricted to intergovernmental cooperation under the principle of non-interference. The Asian financial crisis in 1997, however, provided a strong impetus for East Asian governments to seriously consider formal economic integration within the region. Their desire to strengthen economic and financial cooperation and to

Table 3.2. Inward Foreign Direct Investment Stocks in Europe

		Percent by regional source			Total as % GDP
	Year	Europe	Nafta	Asia	
Germany	1989	58	33	-	7.1
	2000	71	24	-	24.1
France	1989	65	20	-	8.2
	2000	75	21	-	19.9
Italy	1989	78	15	-	5.3
	2000	76	14	-	10.5
UK	1989	40	45	2	20.6
	2000	44	48	3	30.5

Note: The sources are OECD FDI Statistics and UNCTAD International Investment Report.

establish their own regional institutional structure was further spurred by the fear of being marginalised in an increasingly uncertain global trading and financial platform. The failure of the World Trade Organisation (WTO) and Asia Pacific Economic Cooperation (APEC) in making headway on trade liberalisation together with the widening and deepening economic integration in Europe and North America could put individual Asian countries in a significant disadvantage on the global trading platform. Financial surveillance and assistance programs by international supranational organisations had proved to be not infallible during the Asian financial crisis.

There has been much debate about the rise in "Asian regionalism" in recent years and comparisons have often been made between economic integration in Asia and Europe. While we agree with the central view that any emergence of an Asian integration as deep and wide as that of the European Union (EU) is unlikely in the near future, we argue that there are as many economic factors that hold it back as well as the political obstacles that have often been stated as the main impediment towards economic integration. The two major ones are the greater disparity of incomes in East Asia and the pattern of orientation of trade to countries outside the region.

Important similarities exist between Asia and Europe which are both complex and diverse regions with long histories, anchored national cultures and long existing tensions across the different countries. Economically, both regions are open to foreign trade and investment and therefore have a strong common interest in an open international economic and financial system. Serious economic, currency and banking crisis has been experienced by the two regions leading them to share a strong commitment towards the strengthening of economic and financial system so as to mitigate the negative forces of globalisation. At the same time both regions are seeking to preserve the positive features of their complex economic and social models unique to each individual country in the regions.

Notwithstanding these similarities the two regions are also very different. Economic diversity is more pronounced in Asia, ranging from countries with highly industrialised economies to others with traditional, most rural, economic structures. Relative to Asia, EU has a much more homogenous economic grouping even when the poorer candidate countries are taken into account. According to World Bank data, on a purchasing power parity basis, the difference of per capita Gross National Product (GNP) between the richest (Luxembourg) and poorest of the candidate countries (Lithuania) is 650 per cent (see Table 3.3). In Asia, the difference between Japan and Cambodia amounts to about 1900 per cent, but the core countries are not as disparate as

this. In particular the difference between Japan, an industrialised economic power in Asia and China, a demographic and manufacturing giant (700 per cent) is comparable to that between Luxembourg and Romania. The disparities between the 14 larger current members of the EU are much less, and are comparable to those between the four richest East Asian economies. The scale of the disparity between the richest and poorest members of the expanded EU in part indicates the importance of the accession decision made at the Copenhagen summit in late 2002, as they change the nature of the EU from a club of essentially rich countries to a regional organisation designed to promote growth.

A more fundamental difference between the two regions with respect to economic integration is the degree, timing and nature of efforts made towards deeper integration. Driven by stronger shared values created on the rejection of fascism after World War II, Europe has since 1945 been more ambitious than Asia in making explicit a political goal of building an ever-closer union. European integration has proceeded through a single institution based on competition under a common set of rules while in Asia, integration has historically been led by MNCs while recent efforts by governments have created a myriad of forums and competition of initiatives as a consequence of a difference of ambitions between different countries. Furthermore, regional priority has taken centre stage in Europe with reforms imposed for the sake of Europe being more easily accepted and implemented than those decided under the pressure of international organisations and multilateral negotia-

Table 3.3. Gross National Income Per Capita in PPP (2000)

EU	US$	Accession countries	US$	Asia	US$
Luxembourg	45,470	Slovenia	17,310	Japan	27,080
Belgium	27,470	Cyprus	20,780	Hong Kong	25,590
Denmark	27,250	Czech Republic	13,780	Singapore	24,910
Austria	26,330	Hungary	11,990	Korea	17,300
Netherlands	25,850	Slovak Republic	11,040	Malaysia	8,330
Ireland	25,520	Estonia	9,340	Thailand	6,320
Germany	24,920	Poland	9,000	Philippines	4,220
Finland	24,570	Latvia	7,070	China	3,920
France	24,420	Lithuania	6,980	Indonesia	2,830
Sweden	23,970	Bulgaria	5,560	Vietnam	2,000
UK	23,550	Romania	6,360	Laos	1,540
Italy	23,470			Cambodia	1,440
Spain	19,260				
Portugal	16,990				
Greece	16,860				

Note: The source is World Bank.

tions. It is the latter rather than regional negotiations that continue to drive reforms in Asia.

In essence, integration efforts in East Asia are more loosely institutionalised than those in Europe, as is discussed in the Asia Development Outlook (2002). There are no custom unions in the region and no common external tariff structure has been negotiated. Dispute settlement mechanisms tend to be based on bilateral negotiations, with some provisions for resolution of disagreement. Implementation of regulations takes place sub-nationally according to commonly agreed standards. In keeping with its overall philosophy of loose cooperation and open agreements, APEC[3] has few institutional structures. The Asean Free Trade Area[4] (AFTA) was an offshoot of ASEAN which was originally formed for political reasons and need little additional institutional structure to administer. Both APEC and AFTA are aspiring free trade agreements (FTA) between a group of countries to extend reciprocal preferential treatment to the others' exports. AFTA, in particular, has been active in drawing up a timetable whereby tariffs are to be reduced in progressive steps, leading eventually to an FTA.

It has been suggested that Asia could achieve dramatic gains by learning from the experience of European integration which attempts to improve macroeconomic stability and accelerate structural reforms through stable exchange rates, regional macroeconomic surveillance and policy co-ordination mechanism. However, we would argue that many of the gains from European integration arose from its strong internal orientation which became even stronger after the integration.

TRADE PATTERNS IN ASIA

In contrast to the EU, East Asia has historically been externally orientated with its intra-regional trade as a share of total trade in East Asia being significantly lower than that of Europe. The intra-regional export share in East Asia of 33.4 per cent was only slightly more than half of that in EU in 1980. This degree of openness should not come as a surprise given that much of East Asia has so far depended heavily on export led growth which has until recently come from outside of Asia. In particular, East Asia trade has been highly geared towards the US in the aftermath of the second world war due to the dominance of US military power in the Pacific and the fact that it was the only country that emerged from the War with its economy largely intact.

Growing outward direct investments (FDI) from Japan from the late 1980s as well as from the East Asia newly industrialising economies (NIEs)

from the 1990s to the rest of the region has increased the flow of raw materials and intermediate goods within the manufacturing networks in Asia. This coupled with efforts in the late 1990s to strengthened economic integration through AFTA and ASEAN + 3 has increased intra-regional trade somewhat. However the inter-regional export share in Asia increased only by less than 10 percentage point from 1980 to 42.4 per cent in 2000. Asia thus remained significantly less internally orientated than Europe.

We include the five largest East Asia Economies in Table 3.5 in order to make a comparison with Europe. Only in the case of China and Hong Kong is FDI significant. The integration of the region has not gone as far as the creation of single market with limited barriers to the movement of capital, and outward FDI from Japan, for instance is generally in search of cheaper production locations or ones where the barriers in the US and Europe to goods produced by Japanese firms are less stringent. However, we would expect that in the absence of barriers we would see flows of investment from Japan to Korea and the reverse that at least matched the scale seen within Europe in 1989. However, the barriers to such investment are economic, political and cultural, and the process of integration will have to proceed a long way before we see such patterns developing.

The scale of FDI into Hong Kong reflects its decision to put no restrictions on capital movements as well as its role as the major financial and business centre for (southern) China. Much of the FDI in Hong Kong is rebundled and goes on to China, and some FDI from China into Hong Kong returns again to China[5]. In general, much of the inter-Asia FDI has been driven by companies attempting to diversify risk, minimize costs and maximise performance through increased specialisation in order to take full advantage of scale and scope economies. In most cases FDI was welcomed only in the ex-

Table 3.4. East Asian Exports

	Export shares of East Asia (%)	
	1980	2000
East Asia (Intra-regional)	33.4	42.4
US	21.1	23.7
EU	13.4	15.2
Rest of the World	32.1	18.6

Note: The source is Direction of Trade Statistics, IMF

port sector of the host economy while FDI in the domestic markets was subject to numerous restrictions, including an outright prohibition in some sectors. The discriminatory policy of aggressively promoting export-oriented FDI while protecting the local economy from both imports and market-seeking inward investment has inevitably undermined the gains from competition that FDI could bring. Not only has FDI in Asia been highly concentrated in a small number of products e.g. intermediate electronics products, local indigenous companies in host economies had hardly benefited from any spillover such as technology transfer and healthy competition from foreign investors. This pattern of receptivity is clearly different from that in Europe, and competition strategies have been oriented to interactions with the outside world rather than at improving the performance of firms serving the domestic market. Hence the impact on growth will be different.

It is clear from Figure 3.3 on page 58 that the degree of interpenetration even amongst the large economies with Asia is not as advanced as amongst the core Europeans. We should ask ourselves if the level of trade we see within the East Asia block is low or high in comparisons with other regions and the literature on the determinants of trade, and especially of trade gravity models is useful to us here. The basic gravity model has been widely applied as a predictive model for estimating the level of trade between two countries. By analogy to the gravitational force between two masses, the model provides a thorough way to adjust for natural determinants of bilateral trade by assuming that trade between two countries is proportionate to the product of their GNPs and inversely related to the distance between them. Apart from the standard gravity variables, population, physical area,

Table 3.5. Inward Foreign Direct Investment Stocks in East Asia

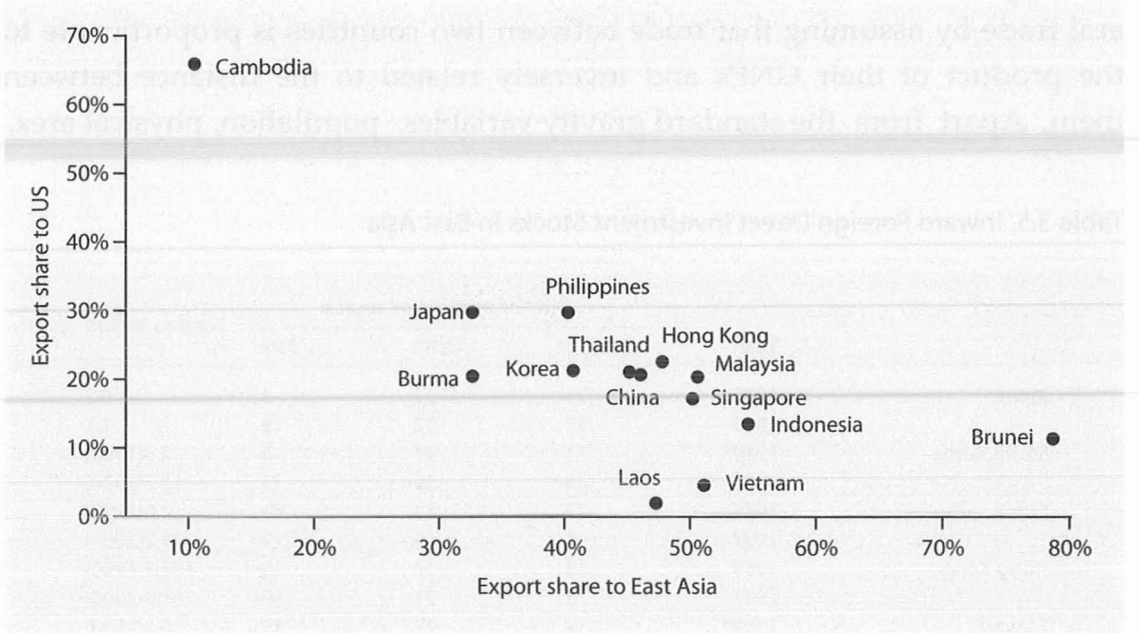
		Percent by regional source			Total as % GDP
	Year	Europe	Nafta	Asia	
Japan	1989	23	52	3	0.3
	2000	52	32	8	1.1
South Korea	1989*	11	27	5	2.3(1990)
	1998	28	35	13	13.7(2000)
China	1989	7	13	66	7.0(1990)
	1997	6	9	72	24.2(2000)
Hong Kong	1989	15	32	15	198.1(1990)
	1997	16	21	19	263.8(2000)
Taiwan	1989*	9	29	19	6.1(1990)
	1999	9	27	20	9.0(2000)

Note: The sources are OECD FDI Statistics and UNCTAD International Investment Report. * indicates that geographical disaggregations of FDI are available only for approved FDI.

common borders and language variables could also be included in the gravity model. In a recent study by the Asian Development Bank (ADB) of 83 countries for the period between 1980 to 2000, it is found that between 68 per cent and 73 per cent of the variance in trade flows could be explained by the basic gravity model. Frankel and Wei (1997) found that while larger and richer economies trade more, trade increases less than proportionally as GNP expands. Bilateral distance has an economically and statistically large negative effect on trade. Common land border, common language and past colonial connection are found to facilitate trade. These factors alone may be enough to explain the higher level of trade integration within Europe as compared to East Asia.

The basic gravity model can be further extended by including dummy variables to capture the effects of economic integration between countries. The impact on trade that could be explained by economic factors common to bilateral trade throughout the world could thus be separated from that coming from regional integration effects. In their study, Frankel and Wei (1997) found that ASEAN countries concentrate their trade among themselves beyond what can be expected from their economic and geographic characteristics. However, this intra-ASEAN bias appeared to have diminished between

Figure 3.3. East Asian Export Shares in 2000



Note: The source is Direction of Trade Statistics, IMF.

1980 and 1992. ASEAN countries are also found to be far more open to the world than an average country. In 1980, the imports by ASEAN countries from the rest of the world were 116 per cent higher than the prediction of the gravity model. This extra-ASEAN openness has further increased by 1992.

A more recent study by Soloaga and Winters (2001) found that intra-ASEAN bias showed a pronounced negative trend between 1987 and 1995 while extra-ASEAN openness was significantly higher in 1995-96 than in 1986-88. While Frankel and Wei (1997) also found intra-EEC bias in Europe, the degree of bias has exhibited a rising trend between 1980 and 1992 in contrast to that in ASEAN. Furthermore, extra-EEC openness was found to have a declined over the sample period. EEC's import from the rest of the world used to be higher than the gravity model prediction by 40 per cent in 1970 but shrank to 26 per cent in 1980 and eventually to zero by the early 1990s. Results from the gravity model study thus lend support to our argument about Europe's internal orientation and Asia's external orientation.

REMOVING CURRENCY BARRIERS TO TRADE

The success of European monetary integration and dissatisfaction with the undesirable aspects of flexible exchange rates has led to renewed discussion of currency pegs in East Asia. A monetary union reduces the barriers to trade even within a common customs area by reducing transactions costs, and this is likely to have a major impact on the level of output that can be produced with a given level of inputs. In addition, a monetary union may increase output because it reduces uncertainty through increasing the coordination of policies and eliminating shocks to bilateral exchange rates.

There is some evidence to suggest that the output gains from the removal of the currency barrier to trade may not be large, reflecting the comparatively weak evidence that exchange rate volatility had a negative effect on trade, but they should be positive. Estimates by Anderton and Skudelny (2001) are representative of the literature they survey. They use a panel for bilateral trade between Euro Area members and six major non-Euro Area trading partners over 1989-99,[6] and find that exchange rate volatility reduced the level of merchandise imports by the Euro Area countries from the non-members by around 10 per cent. Frankel and Wei (1997) using a gravity model argue that few of the intra-regional trade links they study can be attributed to intra-regional currency links. Although real exchange rate volatility has depressed bilateral trade slightly in the 1960s and 1970s, the partially stabilised exchange rates within both Europe and Asia from the 1980s by

countries tracking the Deutsche Mark and US Dollar respectively has had only a small positive effect on trade.

Studies by Rose (2000), Rose (2001), Frankel and Rose (2002) and Glick and Rose (2002) have attempted to assess the impact of monetary unions on trade and output using large panels of countries. They conclude that on average, the level of bilateral trade between two countries with a common currency was significantly higher than might otherwise have been expected after controlling for a number of important factors[7]. Frankel and Rose (2002) also suggest that every 3 percentage point rise in the ratio of international trade to GDP is associated with a rise of 1 per cent in per capita incomes over the following 20 years. These results are controversial, but the debate does suggest that the adoption of common currencies can give significant benefits as long as the accompanying macroeconomic conditions and the choice of partners and pegs are appropriate.

OPTIMAL CURRENCY AREAS AND EAST ASIA

The Asian Financial Crisis in 1997 which exposed the limitations of traditional dollar tracking exchange rate policy used by Asian countries has prompted strong interests in the creation of a collective or cooperative exchange rate arrangement in Asia. The establishment of a Yen-block and an Asian common currency along the lines of the Euro are some of the suggestion that has surfaced in the late 1990s. Williamson (1998) suggested that a natural monetary grouping similar to that of the European Monetary Union (EMU) exists among 9 East Asia economies[8] and argued for the adoption of a common bask peg, surrounded by fluctuation bands of plus-and-minus 10 per cent.

Using optimum currency area (OCA)[9] theory, Eichengreen and Bayoumi (1996) related exchange rate variability to four country characteristics (output disturbances, asymmetric shocks, bilateral trade and economic size) that could influence the desirability of stable exchange rates and monetary unification. When this is applied on Japan and its nineteen leading trading partners over the period 1976–95, they found that OCA does a credible job of explaining the exchange rate policies of Japan's principal trading partners. They further found that the very small and open economies in the Asia region would find it most appealing to peg to other East Asian countries as much benefits could be reaped from the reduction in uncertainty that would result from the creation of a durable common peg. With high levels of intra-Asian trade and investment, quick adjustment to shocks as well as small and

symmetric supply and demand disturbances by European standards, they concluded that the East Asia region satisfied the standard OCA criteria for the adoption of a common monetary policy about as well as Western Europe. Despite this, the study also noted that financial disturbances which are an increasingly prevalent source of asymmetric shocks provide a strong economic case against schemes for a common currency peg especially in Asia where domestic financial systems are less well developed than in Western Europe.

It is clear that many of the shocks that the economy may face do not come through trade. Shocks to consumption can come through common movements in confidence or in asset prices, and a currency area monetary authority needs to respond to these shocks as well. The case for monetary union should in part be based on asking whether shocks are common and in part on asking whether a common response (with no exchange rate variability) is better than separate responses on average over a large number of shocks. The OCA theory was developed in a Keynesian policy making world where demand shocks were assumed to matter, and the Phillips curve was seen as a menu for policy choices. However, developments in theory in economics and the empirical evidence suggest that we cannot increase output by raising inflation, at least in the long run. If there is no long run trade off between output and inflation then there are no long run issues to discuss in this framework. Discussion of policy frameworks in the last decade has moved away from discussing such short-term issues, but they have to be taken into account.

Although there may be a case for fixing exchange rates, the choice of partners is important. Patterns of competition differ from patterns of trade in important ways, and these differences matter when making exchange rate decisions, as the Asia Crisis in 1997 and 1998 made clear. Pegging to a major trading partner, such as the US may be wise if home producers compete with US producers, but it could be problematic if competition is mainly with other non pegged exporters to the US market. In Table 3.6 we look at third party export competition based on the importance of other exporters in the markets that home producers export to.[10] We use 1994 trade patterns as the basis for our calculations as they are more appropriate for our discussion of the impact of pegs during the Asian Crisis.

The large European countries mainly compete with each other in the markets they serve, with indices of over 50%. For them Europe is twice as important a competitor as East Asia in export markets. East Asian producers compete with each other much less in their exports markets than do the Europeans with South Korea as a marginal exception. It would clearly be eas-

ier to form a currency union with a group of trading partners who were also major competitors, as in Europe. If currency peg is forged with a major trading partner as in East Asia before the Crisis when links were made with the USA then there may be major risks. These would be particularly severe when competition was with countries that were not pegged to the dollar, as in East Asia before the crisis. When the dollar euro rate shifted significantly when the Europeans devalued against the USA there were significant implications for exports from East Asia. Although currency pegs were maintained, competitiveness deteriorated sharply in strongly contested markets, and this helped precipitate the crisis. This example makes it clear how important it is to assess patterns of competition in evaluating currency blocks, and the issues differ between Europe and East Asia.

CONCLUSION

European integration has involved three related processes that improve the efficiency of the use of resources. We have seen the removal of tariff and other barriers to trade, the reduction of barriers to competition and the reduction of exchange rate instability and the introduction of a common currency, at least for most members. Each of these steps has had, or can be expected to have a major impact on the level of output. We have seen that increased integration involved increased internal trade as the customs union was perfected, but that increased trade integration has not been as important since 1980. The reduction of exchange rate uncertainty and the introduction of the single market have increased competition and raised the level of investment, and both have had a noticeable impact on output. However, competition can be enhanced both by the movement of firms to markets and by

Table 3.6. Export Competition Index

	Germany	France	Italy	UK
Competition with Europe	0.55	0.56	0.58	0.53
Competition with North America	0.16	0.14	0.14	0.16
Competition with East Asia	0.25	0.24	0.23	0.27
	Japan	South Korea	China	East Asia
Competition with Europe	0.35	0.28	0.57	0.37
Competition with North America	0.26	0.24	0.14	0.26
Competition with East Asia	0.34	0.44	0.25	0.33

Note: The source is NIESR database using 1994 trade weights to compute export competition indices.

the existence of potential contestability. Both have been important in Europe in the last 20 years. Studies of the UK suggest that each of these steps could have (or could raise, if adopted) raised output by 1 to 3 percent.

East Asian countries trade less with each other than did the Europeans in 1960, although the differences are not as large as they were. They also compete much less with each other in their export markets than do the Europeans. Both of these factors would militate against the formation of a common currency in East Asia. However, the gains to increased trade from the existence of a common currency do exist, and they could be a major advantage. In addition a common currency should promote greater real exchange rate stability. This is particularly important for investment and hence for growth, and the macro economic advantages of a common currency could still be large in East Asia, especially as barriers to competition and location are further removed.

Economic integration that goes beyond trade policies, involving cross-border standards and regulations, has also been shown to have important impact on the quality of and commitment to government reform and economic policies. Deep integration necessitates strong policy coordination among countries and peer group pressure has a significant deterrent effect on defection from agreed policy framework as it reduces a country's external credibility. Apart from economic benefits, there are also strong political incentives for economic integration as the trust building effect of increasing integration could reduce the chance of military conflict and raised the level of security in the region. In the post war years, conflict prevention and building of common trust was the vision behind the establishment of the institutions and various policy frameworks that evolved into the current EU. Schiff and Winters (1998) suggests that increased trade among members of a deeply integrated group of countries can increase intra-group trust thereby reducing tensions. Promfet (1996) also suggests that ASEAN may be contributed to the political stability and trust that allowed the formation of cross-border growth zones such as the Singapore-Johor-Riau growth zone, in spite of historical tensions.

The political dimensions driving integration in Europe are important, although the immediate gains appear to be mainly economic. The objective of creating an integrated economic and social space has been central to the very slow and painful process of building an integrated Europe. In order to ensure cohesion in the process of integration the Europeans have adopted an agricultural policy that cannot serve as a model for others, and its costs have offset some of the gains from integration elsewhere. There are other ways to get

the potential gains from competition without the degree of political integration we have seen in Europe, but the adoption of a common legal approach to firms is important in ensuring that relocation can take place. Hence if East Asia were to wish to follow the Europeans without so much political integration then there would still be significant institutional integration. However, some of this would have happened under the aborted FDI agreement put forward by the OECD and the WTO. If we are to re-create an FDI agreement some of the institutional issues on FDI and company legislation will have to be addressed.

REFERENCES

Allen, C., Gasiorek, M. and Smith, A. (1998): 'The competition effects of the Single Market in Europe', *Economic Policy*, 27, 441-486.

Anderton, R. and Skudelny, F. (2001): 'Exchange rate volatility and Euro Area imports', European Central Bank Working Paper No.64.

Artis, M. (2000): 'Should the UK join EMU?' *National Institute Economic Review*, No. 172 January 2000

Barrell, R. and Dury, Karen (2000): 'Choosing the Regime: Macroeconomic effects of UK entry into EMU' *Journal of Common Market Studies*, Vol. 30 (4)

Barrell, R., and te Velde, D. W. (2002): 'European integration and manufactures import demand, an empirical investigation of 10 European countries. *German Economic Review*

Barrell, R., and Pain, N. (1997a): 'European Integration and Foreign Direct Investment: the UK Experience' *National Institute Economic Review* No. 160 April 1997

Barrell, R. and Pain, N. (1997b): 'Foreign direct investment, technological change and economic growth within Europe', *Economic Journal*, 107, 1770-1786.

Dees (1998): 'Foreign Direct Investment in China: Determinants and Effects', *Economics of Planning*, 31(2-3), pp. 175-94.

Durand, M. and Giorno, C. (1987): 'Indicators of international competitiveness: conceptual aspects and evaluation', *OECD Economic Studies*, No. 9.

Durand, M., Simon J. and Webb, C. (1992): 'OECD's Indicators of International Trade and Competitiveness', OECD Economics Department Working Papers, No. 120.

EC (1997): "Trade creation and trade diversion ", *The Single Market Review*, Subseries IV: Volume 3, the European Commission.

Eichengreen B. and Bayoumi, T.: 'Is Asia an Optimum Currency Area? Can It Become One? Regional, Global and Historical Perspectives on Asian Monetary Relations', *Exchange Rate Policies in Asian Emerging countries*, Routledge: London

Frankel, J. and Rose, A. (2002): 'An estimate of the effect of common currencies on trade and income', *Quarterly Journal of Economics*

Frankel, J. and Wei, S. J. (1997): 'Regionalization of World Trade and Currencies: Economic and Politics', *The Regionalisation of the World Economy*, University of Chicago Press: Chicago

Gaisoreck, M., Smith, A., and Venables, A. (2002): 'The benefits to the UK of joining the EU: An Imperfect Competition CGE Analysis' *Journal of Common Market Studies* Vol 23 (4)

Glick, R. and Rose, A. (2002): 'Does a currency union affect trade? The time series evidence', *European Economic Review*, forthcoming, (revised version of NBER Working Paper No.8395).

Hohohan (2001): as a comment on T. Persson, 'Currency unions and trade: how large is the treatment effect?', *Economic Policy*, 33, 435-448.

Pain N. (2002): 'EMU, Investment and Growth: Some Unresolved Issues.' *National Institute Economic Review* No. 180 April 2002

Persson, T. (2001): 'Currency unions and trade: how large is the treatment effect?' *Economic Policy*, 33, 435-448.

Pomfret, R. (1996): 'ASEAN: Always at the Crossroads?', *'Journal of the Asia Pacific Economy'*, 1(3):365-90

Rodríguez, F. and Rodrik, D. (2000): 'Trade policy and economic growth: a skeptic's guide to the cross-national evidence', *NBER Macroeconomics Annual.*

Rodrik, D. (1999): The New Global Economy and Developing Countries: Making Openness Work, Overseas Development Council, Washington, DC.

Rose, A. (2000): 'One money, one market: the effect of common currencies on trade', *Economic Policy*, 30, 9-45.

Rose, A. (2001): 'Currency unions and trade: the effect is large', *Economic Policy*, 33, 449-459.

Schiff, M. and Winters, A. (1998): 'Regional integration as diplomacy', *World Bank Economic Review,* 12(2):271-296

Swann, P., Temple, P. and Shurmer, M. (1996): "Standards and Trade Performance: The UK experience", *The Economic Journal*, 106, 1297-1313

Williamson, J. (1998): 'The case for a common basket peg for East Asian currencies', *Exchange Rate Policies in Asian Emerging countries,* Routledge: London

ENDNOTES

[1] Rodrik (1999) is an intelligent discussion of the interaction of openness and the political process in affecting growth.

[2] Our analysis is based on the current 15 members of the EU and does not include trade with the other members of the effective region, Norway and Switzerland. If we include them regional trade concentration rises to 66 percent in 2000.

[3] Founded in 1989. Members include Australia, Canada, Chile, China, Hong Kong, Indonesia, Japan, Korea, Malaysia, Mexico, New Zealand, Papua New Guinea, Peru, Philippines, Russia, Singapore, Thailand, US and Vietnam.

[4] Founded in 1992. Members include Brunei, Cambodia, Indonesia, Lao, Malaysia, Myanmar, Philippines, Singapore, Thailand and Vietnam.

[5] These issues are discussed at length in Dees (1998).

[6] The six countries are the UK, the US, Japan, Denmark, Sweden and Switzerland.

[7] Estimates vary from their 300% to suggestions by Persson (2001) and Honohan (2001) that the increase in trade could be as high as 40% in the medium term.

[8] These include China, Hong Kong, Indonesia, South Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand.

[9] Artis (2000) discusses OCAs and EMU.

[10] The weight a competitor gets is the sum of their market shares weighted by the importance of the market to the home country. The issues around competitiveness measures is discussed at length in Durand and Giorno (1987) and Durand et al. (1992).

4. International Benchmarking and Economic Performance

Einar Hope*

INTRODUCTION

International benchmarking is high on the political agenda of many countries as a means of evaluating the economic performance of the country in question and as descriptive foundation for formulating policy measures to be taken to improve its economic performance. In Finland two such international benchmarking studies have been undertaken, with some involvement of ETLA in the process.[1] The Norwegian government appointed a public commission in 2000 to develop a system for benchmarking framework conditions for Norway that are relevant for economic growth and value creating activities. I had the privilege to serve as chairman of the Commission. The report was delivered in December 2001.[2] The EU Commission has through the Lisboa process initiated an ambitious and broadly defined benchmarking programme within the EU, where social conditions and policies are also emphasized.[3] Organizations like e.g. the OECD, the UNICE, the World Economic Forum and the IMD are also doing international benchmarking comparisons and analyses.

Benchmarking is a method for measuring and comparing performance, where the outcome typically is related to some measure or identification of best practice. In international benchmarking studies best practice is usually defined in relation to other countries. In the majority of such studies undertaken, there is typically a lot of descriptive information of the country in question and of the specific indicators used for the benchmarking, but surprisingly little discussion in terms of methodological problems and issues with regard to benchmarking. It is surprising, because benchmarking in this context can be easily subjected to methodological criticism on a number of scores. Therefore the question arises whether the often wide-ranging policy conclusions drawn in such benchmarking studies can be sustained or defended on the basis of the underlying analysis, without a thorough discussion of methodological problems and limitations of the analysis.

* Einar Hope is Professor at the Norwegian School of Economics and Business Administration.

In the mandate for the Norwegian Benchmarking Commission (NBC) it was specifically stated that benchmarking as a method should be evaluated and discussed in relation to empirical analysis and policy. The Commission spent considerable time and effort on such an evaluation. I will first review some of the methodological discussion and then give an example of one of the seven specified areas or sectors of the economy, of which the Commission made a specific benchmarking analysis, i.e. education, research and innovation. This area is maybe particularly interesting in a comparison of performance and policy with that of Finland.

BENCHMARKING AS A METHOD

Benchmarking raises methodological questions with regard to the choice of indicators to measure economic performance and their actual empirical measurement, the frame of reference for defining best practice, the comparison between own performance and best practice, and the policy implications to be drawn from the benchmarking exercise. The choice and definition of indicators should have a sound basis in economic theory, but given the state of the art of the theory on the one hand and the empirical and policy orientation of the benchmarking analysis on the other, a balance has to be struck between theoretical prudence and a more heuristic approach. Similarly, a balance has to be struck between benchmarking as a purely descriptive method or tool of analysis and as a normative instrument for policy conclusions. The main steps in a benchmarking analysis can be represented as in Table 4.1.

Benchmarking was originally designed as a method of analysis at the microeconomic level – to compare the performance of a firm or production plant, e.g. in terms of productive efficiency, measured against the best practice either in relation to competitors or to some predefined best practice efficiency standard. Then it was gradually expanded to be applied to cross-country comparisons at more aggregate level, first at the industrial sector level or for a collection of similar enterprises defined on a product or technological basis, and finally to the whole industry of the economy. The latter application is a fairly recent development, though, and raises particular issues of its own.

As seen from the schematic illustration, benchmarking can be used as a descriptive tool to identify and measure strong and weak features of the unit under analysis in terms of indicators of performance. The analysis can end there, as is e.g. the case with the so-called scoreboards or business cycle barometers, just listing the scores on various indicators. This is a rather passive

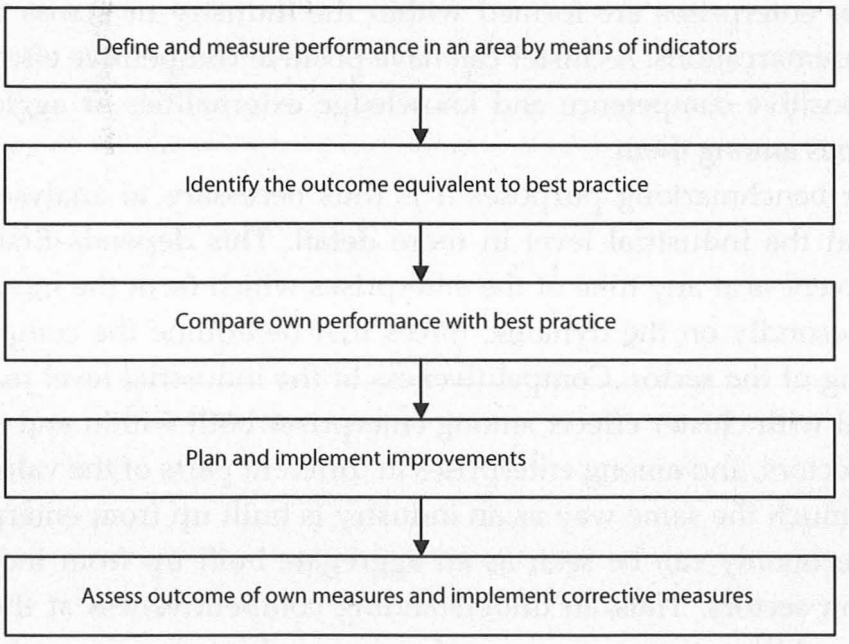
approach to benchmarking, however, and most studies therefore both try to diagnose why the "patient" scores as he does on specific indicators and to suggest remedies to improve the situation if he scores lower than he "should" or could be expected to perform under the circumstances.

COMPETITIVENESS AND BENCHMARKING

Most international benchmarking studies use some measures of national competitiveness as their basic performance indicator. The concept of competitiveness has, however, many dimensions and a number of different definitions can be found in the international literature of the term, e.g. according to level of aggregation. A fundamental question is whether it is possible to define competitiveness in a meaningful way for benchmarking purposes at the aggregate level for the national economy and whether any relevant industrial policy conclusions can be derived from such a benchmarking approach.[4]

It goes almost without saying that an performance indicator like the growth rate of the economy, as used e.g. by the World Economic Forum, is a too aggregate and crude measure of competitiveness to provide much information on the factors influencing competitiveness and on what can be done

Table 4.1. Main Steps in a Benchmarking Analysis



by industrial policy and by industry itself to improve it. In the opinion of the NBC the concept of competitiveness has to be disaggregated and related to the enterprise level, the sectoral level and the national level, respectively, and on that basis prepare indicators reflecting competitive conditions and then comparing the outcome of these indicators to different countries. This is primarily a supply side oriented approach, but demand factors also play a role in the analysis, at least indirectly.

Competitiveness at the enterprise level is derived from the market structure and competitive conditions under which firms have to compete. An enterprise can improve its competitive position by offering its product at a lower price, or a superior quality, or by introducing new products and other innovations compared to the competitors. In order to prepare relevant indicators for benchmarking, one has to analyse the competitive environment at the enterprise level in more detail and describe the strategic competition parameters that firms can exercise in terms of production costs, product prices, quality, product range, accessibility to its products and services, innovations, etc.

An industrial sector is not just a collection of enterprises or plants with similar characteristics in terms of products or technologies, but is also characterized with dimensions like the rate of entry and exit of firms, the restructuring of the industry through mergers or different growth rates of firms of different sizes, the degree of competition among the incumbents and in relation to other sectors, the rate of innovations flowing from the industry, etc. The competitiveness of an industry is also determined by the degree to which clusters of enterprises are formed within the industry or across traditional industry demarcations. A cluster can have positive competitive effects if there e.g. are positive competence and knowledge externalities or agglomeration externalities among them.

For benchmarking purposes it is thus necessary to analyse competitiveness at the industrial level in more detail. This depends firstly on the competitiveness at any time of the enterprises which form the industrial sector and secondly on the dynamic forces that determine the competitive restructuring of the sector. Competitiveness at the industrial level may also be associated with cluster effects among enterprises both within and across traditional sectors, and among enterprises in different parts of the value chain.

In much the same way as an industry is built up from enterprises, the national economy can be seen as an aggregate built up from industries or production sectors. Thus, in understanding competitiveness at the national level, by and large the same approach may be adopted as for understanding competitiveness at the industrial level. A country can in principle maintain its

competitive position if new industries are established at the same rate as old industries decline or disappear, and so that the resources freed by the declining sectors are absorbed by the rise of the new ones.

The competitiveness of the country depends thus on the competitiveness of the enterprises and industries that the economy consists of at any given time and on the dynamic factors that determine the rise and decline of enterprises and industries exposed to international competition. The starting point for such a discussion is, of course, the theory of international trade and comparative advantage, and particularly the modern theory of imperfect competition.[5]

There are some important qualifications to be taken into account, however, when discussing competitiveness at the national level as compared with the enterprise and industry levels. The most obvious one is that enterprises and industries typically consider only the cost of the resources they use in their own production activities, whereas countries must also bear the cost of unused resources, e.g. as a consequence of the restructuring of the economy to maintain its competitiveness. Competitiveness for enterprises, industries and nations depends on the ability to change and adept to new circumstances, but for a country it is also important that its productive resources are fully utilized at any time. If there are rigidities, e.g. in the labour and capital markets, the country may find itself in a position where resources are freed without being transferred to other uses or absorbed costlessly by other sectors.[6] To incorporate such rigidities and find relevant indicators to capture them is an important task in any international benchmarking analysis.

Competition is an inherently dynamic process and should be reflected as such in benchmarking analyses. Thus, performance measures like economic efficiency or competitiveness should be seen as dynamic concepts. In practice one often has to fall back upon static measures, either because of data limitations or because it may be methodologically hard to capture dynamic concepts and processes in operational measures for benchmarking purposes. Take for example the much used indicator in benchmarking analyses of the hourly wage cost for workers in the manufacturing industry. This measure has two severe limitations. Firstly, it is basically static, although if is, of course, possible to estimate changes in it over time. Secondly, the manufacturing industry is only one of many industries that contribute to economic growth and development. At best such measures provide some information about the *competition position*, i.e. the point of departure or starting ground that enterprises, industries and the country as a whole have for competing.

Another dimension of competitiveness to be captured in a dynamic setting in a benchmarking analysis is the capacity of enterprises, industries and economies for change through innovation and restructuring. The NBC places considerable weight on this dimension in its analysis. Innovation is important for dynamic economic efficiency and so is restructuring or adaptability. Innovation is an uncertain process and restructuring is associated with costs. Markets for innovation and knew knowledge do not function efficiently by themselves, so there is scope for an industrial policy to stimulate innovation and knowledge production. Given this ramification, however, in a market economy innovation and restructuring should be based on market signals for profitable investment opportunities and other change parameters. Therefore a fundamental precondition for competitiveness is that markets function efficiently both in a short run and long run perspective.

On this basis the NBC ends up with the following definition of competitiveness:

> "Competitiveness comprises the conditions required for maintaining and increasing economic growth over time at the enterprise level, the industrial level, and the national level. Competitiveness depends on the competition position and the capacity for innovation and restructuring and, at the national level, on securing full utilization of resources and long-term balance of the economy".

DERIVING AND SYSTEMATIZING RELEVANT INDICATORS[7]

On the basis of the separation between enterprise level, industry level, and country level, respectively, and the distinction between the competition position on the one hand and the capacity for innovation and adaptability on the other, a two-way classification scheme can be derived for identifying and systematizing relevant benchmarking indicators. This is illustrated in the Table 4.2.

Factors such as the level and development of labour costs, market shares and realized profitability provide information primarily on competition position and its changes over time, while factors like the rate and direction of innovative activity, the ability to organize and utilize the results of research and development, the efficiency of markets etc are indications more of the capacity of innovation and adaptability. The distinction is not clear-cut, however.

Sometimes a distinction is made between performance indicators and policy indicators, the latter ones being, as the expression indicates, more sus-

ceptible to be operated on from a policy perspective than the first group. As a broad generalization the performance indicators may primarily be associated with the competition position while the policy indicators are more reflective of innovation and adaptability, but again the distinction is not sharp.

The NBC did a pilot-like or preliminary benchmarking analysis, according to the mandate, of seven sectors of the Norwegian economy, applying the methodological approach developed in the study. The sectors are:

1. Natural resources and environment
2. Infrastructure
3. Education, research and development
4. Labour markets
5. Capital markets
6. Product markets
7. Taxation

For each sector a set of main indicators was defined and operationalized that the Commission considered to be relevant and characteristic for the sector under review and that could be used for benchmarking comparisons across countries and over time. The number of proposed indicators for each sector vary from 17 for Labour markets to 5 for Taxation, with a total of some 90 indicators for all the seven sectors. An additional number of indicators are listed as candidates to be included in the benchmarking system when the data and information availability problem can be solved.

Table 4.2. Indicators for Benchmarking at Different Analysis Levels – Some Examples

Analysis level	Competition position	Innovation and adaptability
Enterprise	Profitability	Innovation activity
	Financial strength	Competence
	Market share	Management
Industry	Market growth	Cluster effects
	Establishment rate	Competition
	Internationalization	Knowledge investments
Country	GDP per capita	Education and R&D
	Productivity growth	Mobility of the workforce
	Wage and cost level	Efficiency of capital markets
	Natural resource endowment	Regulatory reforms

On the basis of the indicators the relative strengths and weaknesses of the sectors in terms of competitiveness in relation to the benchmark of other countries were determined. Here I will concentrate on just one of the seven sectors to give an illustration of the benchmarking analysis, i.e. that of Education, research and innovation. In this area 15 indicators were developed and an additional 5 were proposed to be included when data is available.

BENCHMARKING CONDITIONS FOR EDUCATION, RESEARCH AND INNOVATION

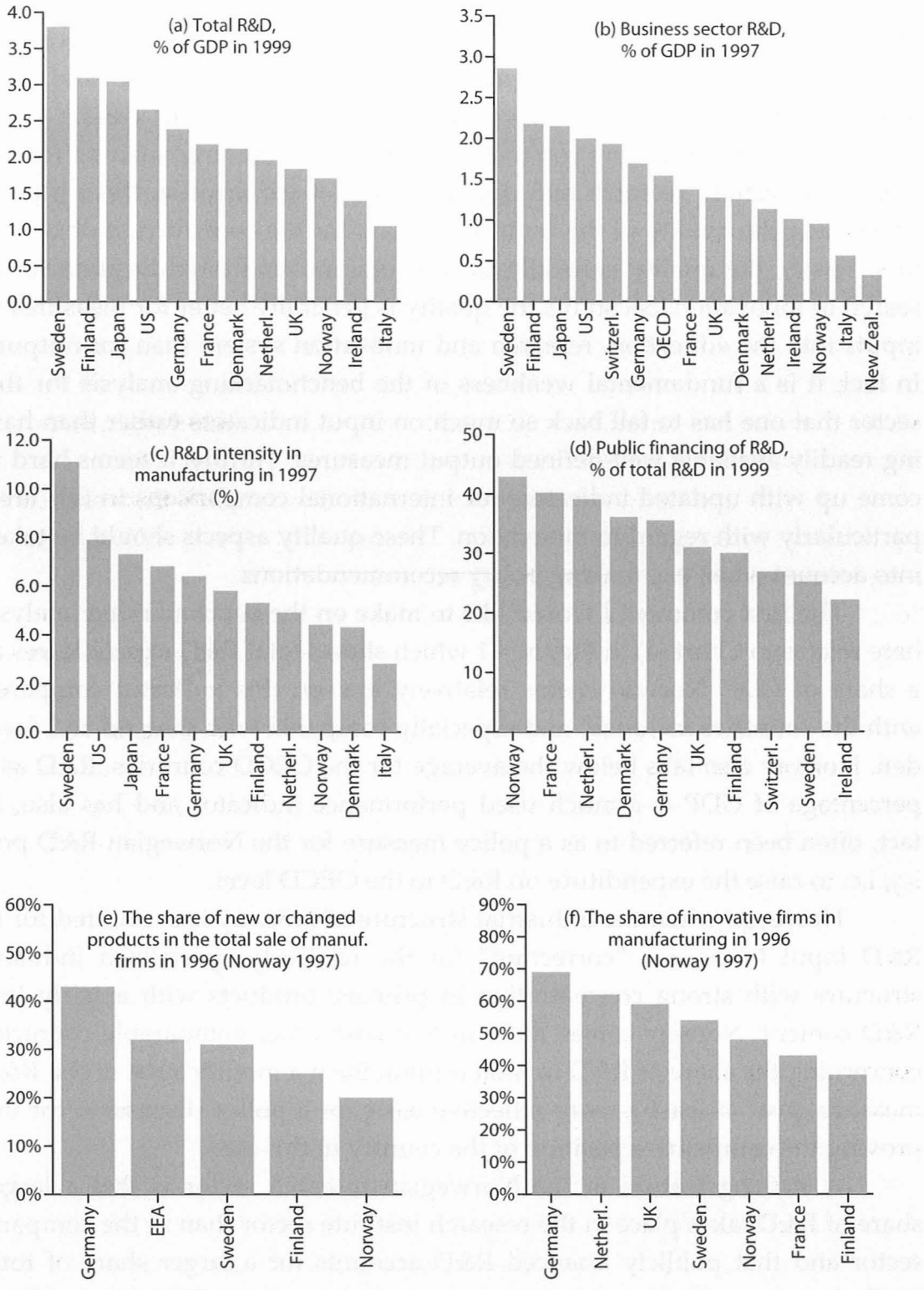
The general, overall characteristics of this sector in an international comparison is given by the NBC as "High level of general education coupled with low investment in research and innovation". This is based on the 15 benchmarking indicators summarised in Table 4.3. Some selected indicators derived for this area by the Commission are illustrated in Figure 4.1.

By aggregating up from individual indicators and on that basis making general qualitative statements in a catch-word manner of the state of competitiveness of Norway with regard to education, research and innovation, the NBC itself may be criticized for falling into the same methodological trap as we so strongly warned against in the methodological discussion, i.e. to draw sweeping conclusions of a policy nature that may not necessarily be sustained by the underlying data. It is a meagre comfort that this seems to be the common procedure in international benchmarking analyses and that one therefore has to abide by the "rules" set by international practice. Personally, as a researcher, I felt somewhat uncomfortable with this stage of the NBC's

Table 4.3. Summary of the Benchmarking Indicators

Strengths
High general level of education
Well developed research institute sector
On par with other countries
R&D within individual industries
R&D per capita
Share of persons with higher education
Challenges
Low R&D in relation to GDP
Low innovative activity and small proportion of new products
Weak interaction between universities and industry

Figure 4.1. Selected Indicators for Benchmarking Conditions for Education, Research and Innovation



Note: The sources are NIFU, OECD (2000) and Nås (2000).

work and would ideally have preferred from this perspective just to have presented the indicators, with due regard to their limitations, and then leaving it to the users of the benchmarking system to make the qualitative generalizations and draw the policy conclusions in a competitiveness context. This said, however, I have no strong qualms with the actual summaries and characterizations that the Commission came up with for the various areas.

Before making some comments on the benchmarking analysis for the area of education, research and development, let me mention three aspects concerning the quality of the underlying data for the construction of indicators. Firstly, the quality generally falls as one moves from education to research to innovation. Secondly, the quality is generally better for measures of inputs into the education, research and innovation system than for outputs. In fact, it is a fundamental weakness of the benchmarking analysis for this sector that one has to fall back so much on input indicators rather than having readily available well-defined output measures. Thirdly, it seems hard to come up with updated indicators for international comparisons in this area, particularly with regard to innovation. These quality aspects should be taken into account when e.g. making policy recommendations.

The first comment I would like to make on the benchmarking analysis here relates to Chart (a) in Figure 4.1 which shows total R&D expenditures as a share of GDP. Norway scores relatively low on this indicator compared with the countries included, and especially compared with Finland and Sweden. Norway also falls below the average for the OECD countries. R&D as a percentage of GDP is a much used performance indicator and has also, in fact, often been referred to as a policy measure for the Norwegian R&D policy, i.e. to raise the expenditure on R&D to the OECD level.

However, when the industrial structure of Norway is accounted for in R&D input terms, i.e. "correcting" for the relatively specialized industry structure with strong concentration in primary products with a fairly low R&D content, Norway comes more in line with other comparable countries concerning the share of R&D in total output. Such a modification of the R&D measure gives at least a more reflective basis for a policy discussion for improving the competitive position of the country in this area.

A striking feature of the Norwegian research sector is that a larger share of R&D takes place in the research institute sector than in the company sector and that publicly financed R&D accounts for a larger share of total R&D than in any other, otherwise comparable country. This can again to some extent be explained by the size distribution structure of Norwegian industrial firms, where the size and structure of the research institute sector

can be seen as a means to compensate for the small size and limited resources for successful R&D performance of the typical Norwegian industrial firm. It is not at all the full story, though.

The NBC classified the well-developed research institute sector as a competitive strength for Norway, but this can, of course, be argued. Whether it is a strength or a weakness depends, e.g., on how well integrated the research institute sector is with the players/firms operating in the market and how well it responds to the demand for knew knowledge and innovation from those players. Many of the research institutes are owned by or operate in close cooperation with the universities and the professional scientific schools. The NBC found a fairly weak interaction between Norwegian universities and industry, and classified this as a weakness from a R&D and innovation perspective. If this weak interaction also spills over to the university connected research institutes, a well-developed research institute sector is not necessarily a strength.

A third "challenge" in a dynamic efficiency and competitiveness context is the relatively low rate of product innovation in the Norwegian economy, as expressed by the proportion of new products in the total product portfolio. This is, of course, related to the low R&D and innovative activity generally, but might also, in the opinion of some commentators, be an expression of a more fundamental lack of creativity, inventiveness and entrepreneurship in the Norwegian society at large. Personally, I am rather sceptical to such a hypothesis or causal "explanation", and it cannot in any case be detected or documented by means of benchmarking, at least not by the type of indicators that are readily available for the international benchmarking of innovative activity with the present data sources.

CONCLUDING REMARKS

Benchmarking is a useful and increasingly used tool of analysis for measuring and comparing economic performance across firms, industrial sectors and countries, but its methodological limitations should be clearly recognized and understood. There are special issues and problems to be considered when comparisons are made at the aggregate national level and when international competitiveness is the performance measure.

The NBC benchmarking study was largely demarcated to the industrial, private sector of the economy. Thus, a comparison of the performance of the public sector was not involved, nor was the importance of a well-functioning public sector for the economic performance of the private sector

considered. The latter question is generally thought of as a matter of economic policy only. It has, however, ramifications for private sector performance well beyond the policy side and should be investigated as such in benchmarking analyses.

An international benchmarking analysis invariably ends up with a discussion of the relevance of economic and industrial policies. Weak economic performance is typically associated with weak policies and thus the need for improved policy performance. This may be so, but it may also contribute to shifting the attention away from the responsibility of the industry itself for its economic performance, independently of the responsibility of the government to further a performance enhancing policy.

Should policy measures be directed at strengthening the competitive strengths further or rather at improving the weaknesses documented in the benchmarking analysis? This is the fundamental policy question. Some policy makers seem to favour a "picking-the-winner" approach which again would largely imply a strengthening of the already strong position approach to policy. Others seem to be more in favour of "let the thousand blossoms blossom" approach, implying a basically neutral policy design in relation to individual firms and sectors. The latter raises a question whether neutrality should be enforced *ex ante* or just *ex post*. In my opinion the principle of policy neutrality should be of overriding importance and that therefore rather strong arguments should be put forward in a given context for deviating from such neutrality, even in an *ex post* context.

REFERENCES

Holmøy, Erling, Hægeland, Torbjørn and Rosendahl, Knut (2002): "Best i test, men hva så? Kommentarer til NOU 2001:29", *Økonomisk Forum*, no 2.

Hope, Einar, ed., (1998): *Competition and trade policies. Coherence or conflict?* Routledge.

International Institute of Management Development (2001): *The world competitiveness yearbook 2001.*

Ministry of Finance (2002): *Finland's Competitiveness and the Way Forward*, Helsinki.

NOU (1996): 17*: I Norge - for tiden? Konsekvenser av økt internasjonalisering for en liten åpen økonomi.*

Nås, Svein Olav (2000): "Innovasjon i Norge - belyst med tilgjengelig statistikk", STEP-rapport R-01.

Nås, Svein Olav (2002): "Innovasjon og FoU: Hva slags resultater kan i forvente å påvise empirisk?", in Einar Hope (ed): *Næringspolitikk for en ny økonomi.* Fagbokforlaget.

OECD (2000): *Main Science and Technology Indicators.*

UNICE (2001): *The ReNEWed Economy. Business for a dynamic Europe. The UNICE benchmarking report 2001.*

World Economic Forum (2000): *The global competitiveness report 2000.*

ENDNOTES

[1] Ministry of Finance of Finland: *Finland's competitiveness and the way forward,* Helsinki, 2002. The report is a sequel to the first Finnish benchmarking study in 1998 by the Ministry of Finance: *Finland as an economic operating environment. Estimation of the strengths and weaknesses of Finnish competitiveness.*

[2] NOU 2001:29. *Best i test? Referansetesting av rammevilkår for verdiskaping i næringslivet.* An English summary has been prepared by the Norwegian Ministry of Trade and Industry: *Best in test? Benchmarking conditions for competitiveness and economic growth*, Oslo, 2002. (www.odin.dep.no/nhd/engelsk).

[3] For benchmarking reports etc by the Commission, see www.benchmarking-in-europe.com/other.html.

[4] Just to cite two definitions; one from NOU 1996:17 *Consequences of increased internationalization for a small, open economy*, p. 8: "The natural definition of national competitiveness in the long term is the ability to hold one's own in international competition while maintaining the remuneration of national input factors at the highest possible level and combining this with full domestic employment. Alternatively, it can be expressed as the country's capacity to ensure the highest possible remuneration of input factors in the society while maintaining full employment and long-term balance in the economy". The other one from World Economic Forum's *Global Competitiveness Report 2000*: "...the set of institutions and economic policies supportive of high rates of economic growth in the medium term". For a critical discussion of industrial policy conclusions to be drawn from benchmarking analyses of international competitiveness, see e.g. comments made to the Norwegian benchmarking report by Holmøy, Hægeland, and Rosendahl in *Økonomisk Forum* no 2, 2002. (In Norwegian).

[5] For a discussion of the relationship between international trade theory and industrial organization theory with regard to analyzing international competition, see the Introduction to Einar Hope (ed): *International trade and competition policies: Coherence or conflict?* Routledge, 1998.

[6] Of particular concern is a situation that often arises under such circumstances, namely that the public sector is used as an instrument to absorb wholly or partly such freed resources, leading to an expansion of that sector in terms of employment. If this development is not reversed after the restructuring is completed, if it ever will be in a dynamic world, this could have a negative effect on the ability of the public sector to serve the private sector efficiently in the long run. This could, again, have a negative influence on the competitiveness of the country.

[7] On the methodological side the NBC did a comprehensive discussion of specific methods and tools for benchmarking analysis, e.g. Salter diagram representations of the production structure in an industry, Data Envelopment Analysis (DEA), enterprise strategy analysis, factor analysis etc., as well as properties of different types of indicator for benchmarking.

5. FDI AND NORWEGIAN INDUSTRY

Per Heum*

For the last 100 years internationally oriented companies have played a significant role in the Norwegian economy. Foreign investors and the industrial competence of foreign multinationals were important to exploit Norway's natural resources, to produce goods that could be sold in international markets, and to lead Norway into a process of industrialization. This preceded Norwegian owned companies, in particular in shipping, which later expanded internationally and became leading in their trades. Over the last decades, also manufacturing companies based in Norway, have become increasingly oriented internationally, while the oil majors have been particularly interested in Norway's oil and gas.

The purpose of this article is to discuss current trends regarding internationally oriented companies and Norway, i.e. I examine foreign direct investments (FDI). FDI increased rapidly on the world scene in the latter decades of last century, however, basically involving industrialized countries. Such investments are considered a major factor in what is referred to as globalisation, as they contribute to economic integration. Thus, data on FDI indicates to what extent the companies of a country are involved in these processes on the world scene, and, furthermore, to what extent internationally oriented investors and companies find it attractive to invest in the country's industrial base. As a natural resource based economy outside the European Union two topics are addressed with reference to Norway. Firstly, the pattern of expansion among the major industrial companies in Norway is considered. Secondly, focus is on inward foreign direct investments, and the attractiveness of the Norwegian economy compared to the economies of it neighbours. In essence, both topics concern the attractiveness for internationally oriented companies to choose Norway as one industrial base for their businesses.

* Per Heum is Director at the Institute for Research in Economics and Business Administration (SNF) in Norway. This article is based on Kåre P. Hagen, Per Heum, Jan I. Haaland, Karen Helene Midelfart and Victor D. Norman *Globalisering, næringslokalisering og økonomisk politic* (Globalisation, location of industry, and economic policy), Fagbokforlaget 2002.

THE INTERNATIONAL ORIENTATION OF NORWAY'S MAJOR MANUFACTURING CORPORATIONS

Internationalisation, or globalisation, means that economic activities are increasingly integrated across countries. To some extent, this is reflected in data on foreign trade and foreign direct investments at the national level. But it is even more evident when considering data on companies. Then, two dimensions are of importance. One is to what extent nationally based companies generate value by having foreign sales, i.e. their value creation is based on sales in foreign market. The other is to what extent the companies base their value creation on production factors abroad, i.e. foreign markets are not served by exports from home based production, but through local production, meaning that the companies have become multinational. The share of employment in the foreign entities of a company may serve as a proxy to indicate the importance of foreign production factors for the value creation of the company.

Table 5.1 shows the share of foreign sales and the share of foreign employment for the 30 largest industrial corporations in Norway in 1980, in 1990 and in 2000. In the small economy of Norway, as in the other Nordic countries, the 30 largest industrial corporations are of great importance for the domestic economy (Heum and Ylä-Anttila, 1993). They make up a significant share of domestic value added in manufacturing, and when it comes to R&D and exports, their role is dominant. Having the focus on to what extent they have sales that are billed abroad, and to what extent their employees work in foreign entities of the companies, we have indicators as to how their value addition activities are distributed between the Norwegian economy and the economies of other countries.

We see from the table that the major manufacturing companies of Norway on the average have had significant foreign sales for decades. However, it is a rather new phenomenon that they are engaged with production

Table 5.1. Sales and Employment of the 30 Largest Manufacturing Firms in Norway (1980, 1990 and 2000)

	1980	1990	2000
Foreign sales, % of aggregate turnover	53	64	78
Foreign employment, % of aggregate employment	9	34	59

Note: The source is SNF.

abroad. As late as 1980, they had more than 90 % of their production in Norway, whereas they had more than half of their sales abroad. Foreign markets were served with traditional exports. Only 20 years later, in 2000, the 30 largest manufacturing corporations of Norway had more employees abroad than in Norway. They have become multinational enterprises, meaning that they increasingly serve foreign markets through local production,

THE MAJOR COMPANIES HAVE INCREASINGLY BECOME MULTINATIONAL

The emergence of multinational enterprises based in Norway, is rather widespread among the 30 largest manufacturing corporations. In Table 5.2 the companies are listed according to how extensively they are engaged in foreign sales and have foreign employment.

It is evident that increasing internationalisation among the largest corporations is not only something that concerns a few. Rather it is a trend which concerns most of them. Today 8 of 10 have more sales abroad than in the domestic economy, compared to about one of three only 20 years ago. Furthermore, today 17 of the 30 largest corporations based in Norway have more employees abroad than in the home country. 20 years ago 7 out of 10 had more than 90% of their employees at home.

Table 5.2. The Distribution of the 30 Largest Manufacturing Corporations According to Foreign Sales and Foreign Employment (1980, 1990 and 2000)

Foreign sales, % of corporate turnover			
	1980	1990	2000
0 - 10%	4	2	3
11 - 30%	5	6	2
31 - 50%	10	4	1
51% -	11	18	24
Foreign employment, % of corporate employment			
	1980	1990	2000
0 - 10%	22	11	6
11 - 30%	6	5	5
31 - 50%	2	8	2
51% -	0	6	17

Note: The source is SNF.

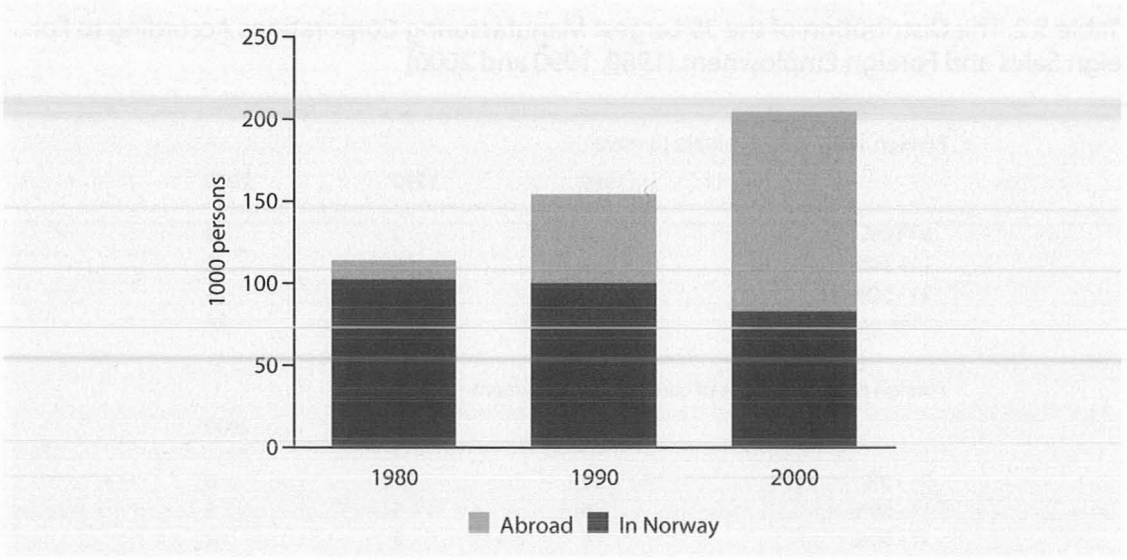
The internationalisation of the largest manufacturing corporations of Norway means that hardly any locate all of their value creating activities to Norway. They have become multinationals, with value creating activities in many different countries. The companies may still be of significant importance to the Norwegian economy, while the Norwegian economy has become less important for their profitability.

GROWTH PATTERN

The internationalisation that we have documented regarding the structure of the major manufacturing corporations of Norway, also show up when we consider the pattern of their expansion. Figure 5.1 presents the number of employees in the 30 largest corporations of 1980, 1990 and 2000, and their distribution between domestic and foreign entities of these companies.

The 30 largest manufacturing corporations in 1990 were significantly larger than the 30 largest in 1980, as the 30 largest of year 2000 also were compared to largest of 1990. Over each of the two decades employment in the 30 largest manufacturing corporations grew by 30% according to employment figures. This growth has to a large extent resulted from mergers and

Figure 5.1. Employment in the 30 Largest Industrial Corporations Based in Norway (1980, 1990 and 2000)



Note: In the figure employment is distributed between entities located in Norway and abroad. The source is SNF.

acquisitions, which have taken place between companies in Norway, and between Norwegian and foreign companies.

It is rather obvious that the growth that follows from this expansion of the largest companies, primarily has taken place outside Norway. Employment in the Norwegian entities of the 30 largest corporations was about the same in 1990 as in 1980, meaning that all employment growth took place abroad. From 1990 to 2000, employment in their Norwegian entities was reduced by 15%, meaning that more than the net growth of the largest corporations took place abroad.

As already illustrated, the structure of the largest corporations is altered, meaning that they no longer are to be considered as national actors. They have become cross-national, or multinational actors, meaning that they are responsible for value creating activities in many different countries. National boundaries, or national jurisdiction, do no limit the scope of their activities as before. Rather they locate the execution of functions and production where they find that the opportunities for profit are the best. The changes in how they have located their value creating activities, seems to indicate that they have found the industrial base of Norway relatively less attractive to the location of activities to the industrial bases of other countries. Whether this ought to concern the Norwegian government depends on to what extent foreign investors also avoid Norway, or not.

NORWAY AS LOCATION FOR INWARD FOREIGN DIRECT INVESTMENTS

The growth pattern of the largest industrial corporations means that they have preferred investments and expansion abroad compared to Norway over the last 20 years. This is chosen for many good reasons, and the growth pattern that is revealed, does not have to be detrimental for the development of the Norwegian economy. After all, the largest corporations of Norway only constitute a minor part of all internationally oriented companies and capital in the world. In principle, Norway is an investment alternative for them all.

Industrial policy should be concerned about how internationally oriented companies and capital in general, whether it is of Norwegian or foreign origin, consider the attractiveness of investing in Norway compared to investments in other countries. It is, however, hard to identify internationally oriented companies and capital in any country, meaning that we are not able to illustrate how internationally oriented Norwegian based investors distribute their investments between Norway and other countries, and how this dis-

tribution has changed over time. Rather, we choose to use the development of inward foreign investments (FDI) as an indicator as to what extent internationally oriented capital chooses Norway as a host country for their investments, and as to the attractiveness of investing in Norway compared to investing in other countries.

Table 5.3 shows that the stock of inward FDI in Norway was at the same level as the EU average in 1990. However, while the share almost doubled in Norway during the 1990s, it almost tripled in the EU, meaning that Norway has become relatively less attractive as an industrial base for internationally oriented investors compared to the EU average.

When compared to the other Nordic countries, it is obvious that Norway is increasingly lagging behind when it comes to the attractiveness of the domestic industrial base for foreign investors. According to the share of inward FDI, Norway used to be the most attractive host for foreign investors up to 1990. However, over the last decade the relative share of inward FDI has increased much more rapidly in all the other Nordic countries. In all of

Table 5.3. Inward Foreign Direct Investments to Norway, the Nordic Countries, and the EU (1990 and 2000)

	1990	2000
Norway	10.7	18.8
EU-15	10.6	30.3
EU-Nordic countries		
Sweden	5.3	36.1
Denmark	6.9	39.6
Finland	3.8	20
EU-Continent, UK and RI		
Netherlands	23.3	65.9
Belgium and Luxemburg	27.8	174
Austria	6.1	16.1
Germany	7.1	68.4
France	8.2	24.1
Great Britain	20.6	19.9
Ireland	7.2	30.5
EU-Mediterranian		
Spain	12.8	25.8
Portugal	14.8	26.5
Italy	5.3	10.5
Greece	9.4	11.1

Note: The values depict the stock of FDI in percent of GDP. Calculations are from World Investment Report 2002.

them they now have reached a level that is higher than in Norway. Also compared to the other EU countries, the attractiveness of the Norwegian industrial base, according to this measure, seems to be lagging behind. In the last decade of the last century Norway was really only outperforming Italy and Greece.

THE ATTRACTIVENESS OF NORWAY'S INDUSTRIAL BASE

To examine why the industrial base of Norway seems to have become less attractive for inward FDI compared to the industrial base of other countries, we need to consider the driving forces of foreign direct investments. In principle, there are three main driving forces, or motives, that may explain FDIs. They have in common that the investors seek access to resources that are rather immobile, meaning that the investors have to establish activities where the resources are in order to access them. We may distinguish between:

- *Factor driven investments,* where companies seek access to productions factors that are localized in specific countries. Natural resources are one example. Another may be low wages.
- *Market driven investments,* where companies access markets which are better served through local establishments. For instance, many services can only be offered when the producer and customer can interact physically. Furthermore, trade barriers, whether they are cultural, political or economical, imply that many markets are hard to serve through traditional exports.
- *Technology driven investments,* where companies seek access to knowledge and technology localized in different countries. When relevant industrial capacity to operate and develop businesses globally are localized in for instance a cluster of companies and research institutions, you have to be part of it to take advantage of its dynamic contributions.

The motives inherent in these driving forces are not mutually exclusive. Nevertheless, in general one will be dominant when it comes to specific investments. For Norway, as the country is richly endowed with natural resources, this means that a significant share of the inward FDIs have been factor driven. Investments have been made to access Norway's natural resources. Today oil and gas attracts the largest interest. Historically, the access to hydro power, and even earlier mining, was the most prominent.

A significant share of inward FDI in Norway has also been marked driven, in the sense that investments are made to access the high purchasing power of the Norwegian inhabitants, or a more sheltered domestic market. This used to be the case, for instance, to serve the national tele monopoly. Today the extensive foreign investments in service production, as in banking, insurance, professional business services, retail trade, hotels and restaurants, all are to be considered as market driven, i.e. to serve demand in these areas in Norway. At least they concern foreign ownership in companies that mainly operate a business that is dedicated to serve this demand locally.

On the other hand, technology driven investments do no seem to play any major role when it comes to inward FDI in Norway. There are examples, however, where the human knowledge base, and the industrial activities in Norway, constitutes an important base for the global businesses of the investor. But by and large, it is not these kinds of investments that characterises inward FDI in Norway. This is also why only a minor share of foreigners' investments in Norway concerns manufacturing.

The rapid growth in foreign direct investments used to be linked to market driven investments. However, throughout the 1990s the growth of FDI globally has increasingly been connected to technology driven investments. Norway's rich endowment of natural resources, and its purchasing power, together with a relatively open attitude to foreign companies and investors, explain why Norway used to be an attractive country for foreign direct investments, for instance compared to the other Nordic countries. When foreign direct investments increasingly are directed towards other European countries, and the other Nordic countries, it is likely to assume that the business and knowledge environment of these countries are more attractive for technology driven investments. Probably the outsider role of Norway with regard to the EU also plays a role, at least for non-EU investors, which want to penetrate Europe. Altogether this raises a lot of challenges for industrial development and future welfare in Norway. The problem for Norway, however, is that these challenges are not likely to appear as urgent for another 10-15 years. This is beyond the time horizon for politicians and the electorate, meaning that Norway most likely will postpone efforts that would have made it easier to handle the return to a non-oil based economy.

References

Heum, P. and Ylä-Anttila, P. (1993): Firm Dynamics in a Nordic Perspective. Large Corporations and Industrial Transformation, SNF-Report No. 78/1993.

World Investment Report (2002): Transnational Corporations and Export Competitiveness, UNCTAD, Geneva.

OSA 2:
PITKÄ AIKAVÄLI – KASVUN LÄHTEET JA SELITYKSET

6. YRITYSKULTTUURIN ROOLI KAPITALISTISESSA TALOUDESSA

Vesa Kanniainen*

ORGANISAATIOT JA KANNUSTIMET

Kansantaloustieteilijät ovat toistaiseksi olleet kohtuullisen rajoitetusti kiinnostuneita yrityskulttuuriin liittyvistä kysymyksistä. Ne ovat perinteisesti olleetkin selkeämmin liiketaloustieteen tutkimuskohde. Tilanne on kuitenkin nopeasti muuttumassa. Kansantaloustieteen piirissä on jo herännyt vahva kiinnostus yritysten hallintakulttuuriin ja omistajavallan rooliin. On käynnistynyt ns. corporate governance -teorian kehittäminen. Sen yksi inspiraatio on ollut yritysorganisaation teoria. Tässä suhteessa on aiheellista mainita Oliver Williamson (1975). Samaan aikaan kansantaloustieteen piirissä onnistuttiin ratkaisemaan kysymys siitä, miten ns. "päämies-agentuurisuhteet" on hyödyllistä formuloida (Holmström 1979). Niinpä esim. liikkeenjohto voidaan tällä kielellä ilmaistuna nähdä yhtiön omistajien (= päämiehet) agenttina. Osapuolten erilainen informaatio heijastuu tällöin heidän välisissä sopimuksissaan. Informaation teorian kehitys käynnistyi Akerlofin (1970) ensin hyljeksitystä ja sittemmin juhlitusta artikkelista. Se tarjosi aikanaan välineet taloussubjektien vuorovaikutuksen analyyttiselle tutkimukselle. Omistajista riippumattoman liikkeenjohdon käyttäytymisen analysointi tuli myös mahdolliseksi.

Keskeinen liitännäisalue yrityskulttuurin tutkimuksen tulemiselle on ollut yritysrahoituksen teoria. Se on osoittautunut haasteelliseksi ja vaatinut monen sukupolven tutkijatyön ennen kuin on syntynyt tunne siitä, että kiinteää maata on jalkojen alla. Kaikki alan tutkijat tiesivät, että Modiglianin ja Millerin vuonna 1958 julkaistu ja sittemmin Nobelin palkintoon johtanut tulos yritysrahoituksen irrelevanssista oli väärä. Tuloksen haastajiksi nousivat kuitenkin vasta v. 1976 Jensen ja Meckling, jotka samaan aikaan taloudellisen agentuuriteorian kehityksen kanssa kehittivät teoriansa rahoituksen agentuurikustannuksista. Heidän lähtökohtansa oli se tosiasia, että yritysjohto tietää enemmän yritystä koskevista asioista kuin ulkopuoliset ja että sillä on

* Vesa Kanniainen on professori Helsingin yliopistossa. Tämän kirjoituksen taustalla olevaa tutkimusta on rahoittanut Liikesivistysrahasto.

kannustin tätä tietoa itsekkääsi hyödyntää. Jensen (1986) eteni tällä tiellä vielä pidemmälle argumentoimalla, että yrityksillä on taipumus kasvaa liian suuriksi, mitä tavoitetta vauhdittaisi johdetun yrityksen resurssien hyödyntämismahdollisuus. Samalla kehitettiin argumentteja, joiden mukaan kapitalismissa tapahtuu virhe-investointeja tai että kannattavia investointeja jätetään tekemättä ja että rahoitusjärjestelmän on niin vaikea tätä ongelmaa korjata (Myers ja Majluf 1984). Vaikka velkarahan käyttö yhtiöissä saa aikaan konkurssiriskin, oivallettiin, että velkaan liittyy hyödyllisenä myös ns. kurivaikutus. Tämä seuraa siitä, että velkasopimuksen velvoitteista laistaminen johtaa yhtiön omistus- ja päätäntäoikeuksien siirtymiseen osakkeenomistajilta velkojille (Stulz 1990). Velkarahan käyttö kuitenkin saattaa aikaansaada moraalisen uhkapelin ongelman, koska se saattaa kannustaa liialliseen riskinottoon. Tämä johtuu siitä, että velkaisen yrityksen osakepääoma on luonteeltaan osto-optio, joka arvo riippuu positiivisesti riskeistä. Esim. pankkikriisien toistuvuus on monasti pyritty selittämään tältä pohjalta. Myös 3G-järjestelmään liittyneiden umts-lupien ylihinnoitellut ja velkarahalla rahoitetut huutokaupat voidaan osaltaan tulkita tästä näkökulmasta.

Pitkään on ollut unohdettuna se tosiasia, että käynnistyvän yrityksen problematiikka on kovin toisenlainen kuin mitä ovat kypsän yrityksen toimintaa luonnehtivat tekijät. Ns. riskirahoituksen noustessa 1990-luvulla ensin USA:ssa ja sittemmin myös Euroopassa merkittäväksi käynnistyvien yritysten rahoitusratkaisuksi oivallettiin, että epävarmuus projektin menestymisestä on suurin projektin alkuvaiheessa (Hall 2002, Geroski 1995), vaikka kvantitatiivisessa mielessä suuryhtiön tappio voi nousta moninkertaisesti. Miten siis on sovittava ns. päätösoikeudet yrityksessä, jolla on riskihanke ja joka tarvitsee ulkopuolista rahoitusta? Mikä on ns. tehokas ratkaisu? Millainen ratkaisu ylipäätään mahdollistaa kannattavien projektien rahoituksen ja toteutumisen. Aghion ja Bolton (1992) esittivät, miten ko. oikeudet tulee kohdentaa projektin menestystä seurattaessa. Koska yrittäjän intressissä on monasti projektin jatkaminen siinäkin tapauksessa, että se näyttää epäonnistuvan ja projekti olisi siksi aiheellista keskeyttää lisätappioiden välttämiseksi, on argumentoitu, että paras ratkaisu rahoituspolitiikassa on velka- ja osakerahoituksen yhdistelmä (Hart 2001). Velan rooli olisi siinä, että sen avulla estyy kassavarojen tuhlaus, mikä vaikutus jo edellä tuli mainituksi ja näin varmistuu paremmin se, että myös ulkopuolisille osakkeenomistajille voi jäädä tuottoa.

Edellä on valotettu yritysrahoituksen roolia yrittäjän ja liikkeenjohdon valvonnassa ja tässä mielessä yrityskulttuurin muovaamisessa. Tervettä riskinottoa ja luovaa tuhoa (Schumpeter 1934) silti tarvitaan innovaatioiden

syntymiseksi ja tuottavuuden nostamiseksi. Osakeyhtiöiden toimintaa kriittisesti arvioitaessa on samalla muistettava, miksi konkurssioikeus, ts. osakkeenomistajien oikeus välttää vastuu yhtiön nimissä otetuista veloista, on yksi toimivan talousjärjestelmän tärkeimpiä innovaatioita maailman taloushistoriassa. Ilman rajoitetun vastuun periaatetta omistajien vastuusta nykymuotoinen suurkapitalismi ei koskaan olisi voinut syntyä (Zingales 2000). Jokainen tuntee idean, että riskien hallinta edellyttää riskien hajauttamista. Tämä johtopäätös pätee kuitenkin vain, kun sijoittajaa suojaa rajoitettu vastuu. Ilman rajoitettua vastuuta kenellekään ei olisi edullista hajauttaa sijoitusta vaan päinvastoin keskittää sijoitus yhteen kohteeseen. Ilman rajoitettua vastuuta riskien hajauttaminen itse asiassa maksimoisi riskit!

Nykykapitalismi on kiistatta tuottanut enemmän vaurautta ja hyvinvointia kuin mikään muu talousjärjestelmä koskaan. Se on tehokkaasti hyödyntänyt aikaisempia innovaatioita ja tuottanut uusia. Huolimatta liikkeenjohdon ja omistajien intressiongelmista yhtiöt palauttavat vuosittain huomattavasti tuottoa sijoittajille. Pitkän aikavälin vertailut osoittavatkin, että osakesijoitusten tuotto on selkeästi ylittänyt velkasijoitusten ja riskittömien sijoitusten tuoton. Kapitalistinen talousjärjestelmä on kuitenkin "keskeneräinen". Yhtiöiden toiminta ja niissä omaksuttu kulttuuri ei aina heijasta riittävää moraalista selkärankaa. Samalla on käynnistynyt keskustelu yhtiöiden yhteiskuntavastuusta.

TALOUSSKANDAALIT

Viime vuosina lehtien taloussivut ovat täyttyneet toinen toistaan räikeämmistä yritysten ja yritysjohdon taloudellisista väärinkäytöksistä, suoranaisista talousrikoksista. Missä on siis talouden moraali?[1]

Liikkeenjohdon optiot ja niiden oikeutus ovat aiheellisesti herättäneet keskustelua kenties enemmän kuin mikään muu teema. On kysytty optiojärjestelmien oikeudenmukaisuuden perään myös valtionyhtiöissä. Optiooikeuksien lunastus ja osakemyynnit kurssihuipulla näyttivät todistavan siitä, että päinvastaisista puheista huolimatta optiot eivät sitouta. Työntekijöiden irtisanomisia on arvosteltu mutta samalla on kyseenalaistettu ay-liikkeen järjestämät tukilakot, joista sivulliset ovat kärsineet ja jotka on toteutettu työehtosopimusten ollessa voimassa.

USA:n tuoreet talousskandaalit ovat olleet tuhoisat ja vauhdittaneet vakavasti maailmantalouden tilan heikentymistä.[2] Niinikään kirjanpitotoimistojen, analyytikkojen ja investointipankkiirien moraali on aika ajoin laajalti jouduttu kyseenalaistamaan.

Talous ilmeisesti tarvitsee eettisen suihkun. Tarvitaan kapitalismin kriittistä arviointia, omistajavallan analyysia ja tehostamista mutta myös kapitalismin puolustamista siinä, missä se on saanut aikaan hyvää.

- Mikä selittää markkinavoimien moraalikadon ongelman? Olisiko se vältettävissä tai hallittavissa? Miten yrityskulttuuria koskeva teoria voi auttaa ymmärtämään ja selittämään talousskandaaleja?
- Miten kurittaa kapitalismin pettureita ja estää uusien nouseminen? Onko talousmaailman sankarin kaavun alla sittenkin aina petturi? Mitä on ahneus? Mikä on oikein, mikä on väärin?[3]

Tuoreissa skandaaleissa omaa etuaan ovat olleet tavoittelemassa kapitalismin jalokivien vartijat ja talouselämän juhlitut sankarit. Ihailluissa suuryrityksissä paljastui suoranaista talousrikollisuutta ja suuren luokan väärinkäytöksiä. Pettymys tuli siitä, että skandaalit olivat mahdollisia niin suurimittaisina. Yritysjohdon ahneutta olivat ruokkineet satumaiset hyötymismahdollisuudet. Kiinnijäämisen, maineen menetyksen, sakkojen ja vankilareissun jälkeenkin käteen jäi potti, joka oli niin mittava, että keikka kannatti ja riski kannatti ottaa. Kapitalismi tuntui kääntävän yleisöön päin rumat kasvonsa, kun häikäilemätön oman edun tavoittelu, suuren luokan huijaukset ja petokset alkoivat paljastua isossa mittakaavassa etenkin USA:ssa.

Kapitalismin menestys edellyttää vahvaa sosiaalista pääomaa, luottamuspääomaa. Mittavia liikesopimuksia ja yhteistyökuvioita voidaan tällöin saada aikaan mobilisoimatta samalla lakimiesten armeijaa. Yhteisistä pelisäännöistä luominen ei tarkoita sitä, että asiat aina viedään tuomioistuimiin. Yritysten verkottuessa liiketoimet perustuvat yhä enemmän luottamuspääomaan ilman oikeudellisia kytkentöjä. Kapitalismin toimivuuden kannalta on kuitenkin tärkeätä, että taustalla on oikeusvaltio oikeusjärjestelmineen. Siihen voidaan vedota, jos joku rikkoo pelisääntöjä.

YRITYSTOIMINNASTA

Skandaaleista huolimatta onkin tunnustettava, että juuri kapitalistinen talousjärjestelmä on luonut tuloa, vaurautta ja materiaalista hyvinvointia tehokkaammin kuin mikään muu järjestelmä. Kapitalismi voitti ideologioiden taistelun, koska vain kapitalismi on onnistunut luomaan yhteistä taloudellista hyvää jaettavaksi asti, mikä puolestaan on esimerkiksi ehkäissyt räikeää korruptiota. Tutkimustyö on auttanut ymmärtämään, miksi tämä on ollut mahdollista. Ja se auttaa myös ymmärtämään, miksi moraalikatoa voi syntyä.

Ihmisten haluun tulla onnelliseksi ja elää turvattua elämää mikään ihmisen järjestelmä ei silti näytä vastanneen. Maailma on riskeistä rakennettu, niin taloudellisista riskeistä kuin muistakin uhkista. Riskit eivät olekaan kapitalismin syytä.

Hyvinvoinnin ja elintason luomiseen tarvitaankin sekä järjestäytynyttä yhteiskuntaa että yksityistä yritystoimintaa. Näin syntyy arvonlisää eli palkkaa työntekijöille ja voittoa omistajille. Se on sitä arvonlisää, jota julkinen valta verottaa ja jakaa tulonsiirtoina ja peruspalveluina kansalaisille.

Järjestäytynyt yhteiskunta edellyttää vakaata julkista sektoria, joka ottaa vastuun sisäisen turvallisuuden ja muiden ns. julkishyödykkeiden tuottamisesta. Ilman riittävää sosiaalista pääomaa, luottamuspääomaa ja yhteiskunnallista infrastruktuuria kehitys ei voi edes käynnistyä. Toiseksi, tarvitaan yrityssektori, joka luo taloudellisen arvonlisän, löytää kannattavat investointikohteet, tuottaa riittävästi tuotantoa ja voittoa. Vain tämä mahdollistaa työllistymisen ja tulonmuodostuksen, jota voidaan jakaa ihmisille ja verottaa yhteiseen kassaan. Kaikki tämä edellyttää myös tarpeellisten instituutioiden, kuten rahoitusjärjestelmän, sekä taloudellisten ja ei-taloudellisten kannustimien kehitystä ja rakentamista. Kansainvälinen vaihdanta saa aikaan paljon hyvää. Sen estäminen kuitenkin on niin monen intressissä ja kehityksen este.

Yritystoiminnan kannustinjärjestelmien on oltava oikeita. Ei ole kuultu kenenkään arvostelevan Nokian johdon optioita eikä Nokian voittoa ylisuureksi. Jos Nokia olisi samassa tappiokierteessä kuin Ericsson, sen pääkonttori tuskin sijaitsisi Suomessa ja sen suomalainen johto olisi jo korvattu ei-suomalaisella miehityksellä.

Yritystoiminnalla on liian monen suomalaisen korvissa edelleen vahva negatiivinen kaiku. 60-luvun laulut soivat vieläkin korvissa, vaikka uusi aika vaatisi uuden sävellajin. On silti paljolti yrittäjien ja liike-elämän oma vika, jos yritysimago on kielteinen. Yksi skandaali tai talousrikos heittää helposti varjon koko liike-elämän ylle.

Suomen traumaattinen lama opetti, kuinka tärkeätä on, että yritykset voivat hyvin. Työpaikat ja arvonlisä luodaan yrityksissä. Yrittäjiksi, riskinkantajiksi, hakeutuu työikäisestä väestöstä vain runsas 7 %. Rehellinen yrittäjä on kansantalouden hyväntekijä. Kielteisten uutisten alle hautautuu tosiasia, että hyvinvointi perustuu yritysten luomaan arvonlisään.

Pääomatulojen osuus vahvistui korkeasuhdanteessa 1990-luvun jälkipuoliskolla. Suhdanteiden kääntyminen on kuitenkin jo paljastanut taloudelliset riskit. Ne näkyvät ensimmäiseksi yritysten tuloksissa. Tämä näkyy myös osakkeiden kursseissa, jotka heiluvat päivittäin. Hyvinvointi on uhattuna.

USA:n piilaakson teknoyrityksistä on kaatunut kolmannes ja muut voivat huonosti.[4] BKT:n tuottaminen ja ylläpitäminen vuosittain edes entisellä tasolla on pieni ihme. Maailmantalous rypee tätä kirjoitettaessa vaikeuksissa: Japani, Saksa, Etelä-Amerikan suuret maat, USA. Kapitalisoitunut ja globalisoitunut Kiina näyttää olevan lähes ainoa merkittävä eteenpäin menijä.

Ilmeisesti maailmaa voidaan parantaa vain, jos ymmärretään ne lait, jotka talouselämää säätelevät. Ne ovat kirjoittamattomia lakeja. Niiden löytäminen on kansantaloutta koskevan tutkimustyön tehtävä. Eikä niiden tarjoamaa ymmärtämystä voi korvata mutu-ajattelu.

Kaduilla mellastavan nuorison tuskan ymmärtää. Historia kuitenkin todistaa, että maailman tekeminen paremmaksi epäonnistuu ja tuo kurjuutta, jos ihmisyyttä ja maailman kehitystä hallitsevia kirjoittamattomia lakeja ei haluta ymmärtää.

KAPITALISMIN HOUKUTUKSET

Tuoreiden havaintojen perusteella näyttää korostuvan käsitys, että suuryritysten johto mitä ilmeisimmin painottaa taloudellisissa päätöksissään lyhyen aikavälin hyötyjä monestakin eri syystä. Vuositilinpäätöksistä on siirrytty kvartaalikapitalismiin. Samalla optiojärjestelmien myötä henkilökohtaisten etujen tavoittelun mahdollisuus on vahvistunut. Voi myös olla, että johto pelkää osakkeen aliarvostusta markkinoilla ja siten johtamansa yrityksen vihamielistä yritysostoa. Kone osti Partekin, ja johto sai lähteä. Alihinnoiteltu Instrumentarium ostettiin USA:han.

Johdon tiedossa oleva katastrofiskenaario saattaa johtaa houkutukseen manipuloida osakekurssia lyhytnäköisesti ts. jättämällä kannattavia hankkeita toteuttamatta, yliarvostamalla yrityksen omaisuutta, salaamalla markkinoilta riskillisiä operaatioita, joita yritys on toteuttamassa, salaamalla johdon tiedossa olevia luottotappioita jne. (Stein 1989). Kyse on tuloksen vääristämisestä ja sidosryhmien harhauttamisesta. Syntyy moraalikadon ongelma. Syntyy kiusaus, mahdollisuus opportunismiin.

Näkyvissä oleva katastrofi on toisinaan johtanut itsemurhaan. Mahdollisuus tiedon manipulointiin voi johtaa myös Enron-ilmiöön, jossa rikollisesti vääristellään tulos. Johto vie kaiken, mitä on saatavilla, ja osakkeenomistajat ja sidosryhmät häviävät.

Ihminen näyttää olevan perimmältään opportunistinen ja itsekäs eettisissä valinnoissaan. Onneksi tämä ei ole kuitenkaan koko totuus ihmisestä ja ihmisyydestä. Vaikka Raamattu onkin oikeassa ihmisen raadollisuudesta,

kokeelliset tutkimukset kuitenkin viestittävät myös siitä, että ihmiset arvostavat oikeudenmukaisuutta, luottamusta ja vastavuoroisuutta. Meillä on siis sittenkin omatunto.

Silti opportunismi ja oman hyötymisen tavoittelu saavat ajoittain voiton eivätkä tietenkään rajoitu kapitalismiin. Kiusaus on järjestelmästä riippumaton. Demokraattisesti hallitut kapitalistiset yhteiskunnat sitä paitsi ovat kehittäneet muita tehokkaampia keinoja yrittää paljastaa ja tuomita huijarit.

YRITYKSILLE KUULUU YHTEISKUNTAVASTUUTA

Friedmanin (1970) mukaan yrityksen ainoa vastuu on voiton tekeminen. On kuitenkin kysytty, kuuluuko yrityksille muutakin vastuuta. Yrityksillä on etuoikeus toimia ja tehdä tulosta järjestäytyneessä yhteiskunnassa. Ne saavat nojata toimissaan sen arvokkaaseen sosiaaliseen pääomaan ja luottamukseen. Ne saavat hyödyntää aikaisempien sukupolvien luomaa infrastruktuurin ja muita julkishyödykkeitä. Ja ne saavat ostaa työmarkkinoilta yhteiskunnan varoin koulutettua, osaavaa työvoimaa.

Siksi yrityksenkin vastuulla on viestikapulan vieminen eteenpäin sosiaalista pääomaa vahvistamalla, eettisesti vahvojen arvojen edistämisellä. Näin syntyy tuottoa yhteiselle sosiaaliselle sopimukselle. Näin vahvistuu normisto, sosiaalinen omatunto.

Yhteiskuntavastuu voi muuttua myös kilpailueduksi. Kuluttajat ovat valmiit maksamaan laadusta (Shapiro 1983). Kansalaisilla ei silti voi olla kohtuuttomia odotuksia yritysten yleisvastuusta, eikä yritystenkään ole aihetta yrittää muuttua hyväntekeväisyyslaitoksiksi. Niiden päätehtävä on tehdä tulosta mutta tehdä tulos vahvojen arvojen keinoin. Friedman on oikeassa mutta vain osaksi. Sidosryhmävastuun noustua uudelleen keskusteltavaksi on toisaalta unohdettu maailman muutos. Aikanaan yrityksillä oli tapana huolehtia esim. työvoimansa asunto-oloista, koulutuksesta ja muutoinkin työntekijästään. Hyvinvointiyhteiskunnan kehityttyä tämä vastuu on luontevasti siirtyneet yhteiskunnalle.

MITKÄ TEKIJÄT SÄÄTELEVÄT YRITYSKULTTUURIN MUODOSTUMISTA?

Yrityskulttuuria on vaikeata mitata.[5] Tämä ongelma ei kuitenkaan tee yrityskulttuurista vähemmän tärkeätä tutkimuskohdetta. Se tekee siitä pikemminkin haastavamman. On nimittäin ilmiselvää, että yrityksen sisällä kehittyvä kulttuuri voi vaikuttaa ratkaisevastikin yrityksen menestymiseen markkinoil-

laan ja siten kannattavuuteen. On myös havaittu, että eri maissa voi olla kovin toisistaan poikkeava yrityskulttuuri. Verrataanpa vaikka japanilaisia, länsi-eurooppalaisia, venäläisiä ja amerikkalaisia yrityksiä.

Taloustieteessä on jo tutkittu yritysten hallintajärjestelmiä (corporate governance). Pääpaino on sen kysymyksen valaisemisessa, miten yrityksen omistajien etu näkyisi yritysjohdon toiminnassa. Voi ajatella, että yrityksen sisällä muodostuva yrityskulttuuri osaltaan toimii hallintajärjestelmänä, ulkoisen kontrollin substituuttina (Thomsen 2001).

Yrityskulttuurin talousteoria on suppea, mutta itse asiassa jo olemassa alkaen Krepsin (1990) artikkelilla. Erinomaisen syvällisen ja kokoavan katsauksen on sittemmin esittänyt Hermalin (2002). Taloustieteen suhteellinen etu löytyy tälläkin alalla sen metodisesta ylivertaisuudesta muihin tieteenalueisiin nähden. Voidaan viitata peliteoriaan, informaation teoriaan ja organisaation teoriaan. Yritysten sisällä eri osastoilla tai hierarkiatasoilla toimivilla voi olla omat tavoitteensa, jotka eivät ole muiden kanssa yhteneviä. Yhteistyö on kuitenkin tärkeää. Periaatteessa sitä voidaan vahvistaa muodollisten sopimusten avulla. Kreps (1990) oivalsi, että muodolliset sopimukset ovat hankalia neuvotella ja valvoa. Vuorovaikutuksen toistuvuus ja pitkäaikaisuus ovat yhteistyökykyä vahvistava mekanismi, joka ei tarvitse muodollisia sopimuksia tuekseen.[6]

Intuitiivisesti voi sanoa, että yrityskulttuuri "tänään" on jotain, joka perittiin tavasta toimia "eilen". Kulttuurissa on jotain pysyvää. Se merkitsee ennustettavuutta yrityksen päätöksenteossa etenkin suhteessa henkilöstöpolitiikkaan. Vastuussa päätöksistä oleviin liitetään heidän maineensa siitä, miten he ovat toimineet aikaisemmin. Tämä vahvistaa ennustettavuutta, millä hyvinkin voi olla myönteinen kannustinvaikutus yrityksen tulokseen.

Kysymys ei silti välttämättä ole tavoitekonflikteista yritysten sisällä. Voi myös olla, että eri ryhmillä on poikkeava kyky vastaanottaa, prosessoida ja välittää informaatiota. Kommunikaation tehokas organisoiminen voi siten olla yksi yrityskulttuurin tehtäviä. Tästä näkökulmasta yrityskulttuurin kehittäminen vastaamaan tätä tehtävää voi toimia otollisena investointikohteena (Cremer 1993). Lazear (1995) on analysoinut yleisestä näkökulmasta kulttuurin sisäistymistä ryhmässä evolutionaarisen teorian avulla. Hermalin (2000) on antanut tälle teorialle suoraviivaisemman ja yrityskulttuuria koskettelevan tulkinnan. Ideana on arvioida, miten organisaatiossa kehittyvät sellaiset preferenssit, joita liikkeenjohto haluaa vahvistaa. Niinpä vahvistaminen voi tapahtua "stalinistisesti" eli vaihtamalla ihmisiä tai "maolaisesti" eli pyrkimällä heitä muuttamaan uudelleen kouluttamalla.

Itsekkään egoistin talousteoria on viime vuosina tullut haastetuksi, kun on havaittu, että ihmiset laajalti käyttäytyvät altruistisen vastavuoroisuuden hengessä. Ihmiset rankaisevat poikkeavaa käyttäytymistä, vastaavat saamaansa lahjaan lahjalla ja käyttäytyvät reilusti myös tilanteissa, joissa siitä ei taatusti palkita (juomarahan jättäminen). Evoluutiobiologiassa tällainen käyttäytyminen on ollut kauan tunnettu (Wilson 1975). Oivan katsauksen on esittänyt Ridley (1996). Taloustieteen puolella Frank (1998) on analysoinut tunteita ja luottamusta. Kandel ja Lazear (1992) ovat analysoineet syyllisyyden ja häpeän merkitystä inhimilliselle käyttäytymiselle. Hermalin (2000) on osoittanut, miten tämä voidaan formuloida Holmströmin (1982) tiimi -mallin avulla. Taloustieteilijän vastaus johtajuuteen ilmeisesti on toinen kuin mitä muut tieteenalat tarjoavat. Taloustieteilijälle johtajuus perustuu informaatioetuun (Hermalin 1999). Yrityskulttuurin analyysissa myös delegoinnin teoria on tärkeä. Kaikki nämä suuntaukset osoittavat, että talousteorian mahdollisuus tukea vahvana metoditieteenä muita käyttäytymistieteitä avaa aivan uusia mahdollisuuksia ymmärtää entistä paremmin myös yrityskulttuureita ja niissä vallitsevia lainalaisuuksia.

Eri toimialoilla tai tuotantolaitoksissa saattaa yksilösuoritusten ja ryhmäsuoritusten tarve olla erilainen. Hyvä työilmapiiri ilmeisesti on sitä tärkeämpi, mitä arvokkaampi on työntekijöiden keskinäinen yhteistyö yrityksen menestymisen kannalta, ts. mitä merkityksellisempi on se, että työyhteisö muodostaa tiimin. Yrityksen on kuitenkin vaikeata palkita työntekijäänsä yhteistyökykyisyydestä. Kuten Rob ja Zemsky (2002) ovat havainneet, tämä saattaa kuitenkin olla epäsuorasti mahdollista vahvistamalla työntekijöiden henkilökohtaista hyvinvointia yrityksen sisällä vallitsevan "sosiaalisen pääoman" vahvistamisen kautta.

Tayloristiseksi kutsutaan organisaatiota, jossa vallitsee vahva spesialisoituminen tehtävien mukaan. Holistisissa organisaatioissa taas tehtävien ja työvoiman "matching" on joustavampi. Lindbeck ja Snower (2000) ovat esimerkki analyysista, jossa työvoimaa käytetään komplementaarisesti eri tehtäviin ja saavutetaan tehokkuusetuja. Työnjakoa yrityksen sisällä voidaan toteuttaa sentraloidusti tai hajautetusti osaamisen syntyessä tekemisen kautta (Greenan ja Guellec 1994).

MITEN KONTROLLOIDA LIIKKEENJOHTOA?

Liikkeenjohdon moraalia arvioitaessa on yhtäältä oltu huolissaan siitä, että omistajista riippumaton johto väärinkäyttää yrityksen varoja. Pahimmillaan johto voi varastaa omistajien varallisuutta ja tuhota myös työntekijöiden elä-

kesäästöt. Yritysjohdon optiot ovatkin ruokkineet kulttuuria, jossa johtajat ovat yrittäneet nostattaa osakkeen arvoa kaikin keinoin, koska ovat itse hyötyneet siitä rahallisesti. Talousrikokset ovat luovaa toimintaa, luovuuden ja lahjakkuuden hedelmiä nekin. Niistä seuraa kuitenkin luovaa tuhoa kielteisessä mielessä.

Toisaalta taas on esitetty vastakkainen huoli ts. huoli siitä, että vahvan omistajavallan ollessa toimiva ja johdon toimiessa osakkeenomistajien etua ajaen yritysten muut sidosryhmät (työntekijät, lainanantajat, tavaratoimittajat jne.) kärsivät.

Kapitalismissa toki tapahtuu valtaisa määrä liiketoimia päivittäin ilman kirjoitettuja sopimuksia ja ilman ongelmia. Huolimatta paljastuneista huijauksista sijoittajat ja säästäjät ovat saaneet huomattavan määrän tuottoa niille säästöilleen, jotka he ovat luovuttaneet liikkeenjohdon käyttöön. Tuoreet skandaalit johtavat silti tarpeeseen uudelleen arvioida yritysten ja liikkeenjohdon toimintaa. Yritysten ja markkinavoimien täysi vapaus sisältää kriisien ainekset. Tarvitaan järjestelmiä, jotka edesauttavat luottamuspääoman vahvistumista. Mitä ne voivat olla?

Kirjanpidon oikeaksi vannominen ja rikosoikeudelliset seuraukset ovat sopiva pelote, mutta tuskin ratkaisu. Keinoina voivat sen sijaan olla eettiset tilinpäätökset, omistajavallan tehostuva käyttö ja sisäpiirimääräysten kiristäminen.

Eettisten tilinpäätösten tavoittelu on kaunis ajatus. Se perustuu vanhaan viisauteen siitä, että "Hyvä kello kauas kuuluu, paha vielä kauemmas". Luotu imago ei kuitenkaan saisi poiketa todellisesta. Ei siksi ole aihetta olla liian optimistinen sen suhteen, että julistukset eettisestä arvomaailmasta yksinään voisivat olla ratkaisu. Myös Enronia pidettiin vahvojen arvojen yrityksenä. Markkinoiden kyky rangaista ajoissa on rajallinen.

Sisäpiiriongelma on sitä, että yhtiön sisäpiiriläinen hyödyntää omaksi edukseen tietoa, joka sillä on ennen muita tahoja. Oikeampi periaate olisi, että johto ottaisi omansa vasta, kun pörssitiedotteet on annettu. Muutoin voi syntyä moraalikadon ongelma. Sijoittajien harhauttaminen on helppoa. Nämä kun käyttäytyvät kuin laumaeläin ja menevät ryhmän mukana.[7]

Tulisiko siis sisäpiiriläiseltä hänen suunnitellessaan osakeomistuksensa myyntiä edellyttää muutaman viikon ennakkovaroitus, jotta markkinat tietävät määrätä oikean hinnan? Näin vältettäisiin yllätyksellinen oman yrityksen osakeomistuksesta irtaantuminen.

Entä omistajavalta? Omistajavallan tehtävänä on löytää tasapaino kahden eri tavoitteen välille, ts. on taattava liikkeenjohdolle riittävä vapaus tehdä kauaskantoisia päätöksiä, mutta samalla on pyrittävä estämään väärin-

käytökset. Johdon palkkaus tehtiin riippuvaiseksi tuloksesta (Grant, King ja Polar 1996). Ajateltiin, että markkinat kurittaisivat.

Omistajavallan yksinkertaistettuna ideana on tuottaa päätöksiä, jotka maksimoivat voittoa ja siten osakkeen markkina-arvoa. Tällä teorialla on laajalti kovin huono kaiku. Yhdistyneenä onnettomasti rakennettuihin johdon optioihin se on myös ollut johtamassa skandaalimaiseen mahalaskuun samalla, kun johdon huipputulot ovat karanneet käsistä. Omistajakeskeisestä ajattelusta ja optioista luopuminen ei liene silti edessä. Tutkimus ei nimittäin tue sitä kaunista ajatusta, että tavoite sidosryhmien hyvinvoinnista huolehtimisesta laajemmin voisi sittenkään olla parempi tavoite liikkeenjohdolle määrättäväksi.

Pahimmillaan sidosryhmäpainotteisuus voi johtaa todellisiin katastrofeihin. Kaikki huonotkin liikkeenjohdon ratkaisut olisivat perusteltavissa sillä, että ne ovat ainakin jonkin sidosryhmän etujen mukaisia. Sitä paitsi pörssissä on varoittavia esimerkkejä. Omistajavallan korvaaminen jollain muulla vallankäytöllä olisi siksi ongelmallinen (Tirole 2001). Näin on maailmamme valitettavasti rakennettu.

Taloudellisten instituutioiden luominen ja kokeileminen ovat kuitenkin darwinistinen prosessi, jossa paremmin toimivat ratkaisut korvaavat ajan myötä heikommat. Tätä kehitysoptimismia tukee koko taloushistoria: innovaatiot, rahoitusjärjestelmän kehitys, osakeyhtiömuodon kehitys jne. Julkista regulaatiota tarvitaan mutta niin, että se edesauttaa oikeiden kannustimien vahvistumista.

On ilmeisesti hyvä, että päätösvalta on tahoilla, joiden vaateet yrityksen tulokseen edustavat ns. residuaalioikeutta eli tulevat nokkimisjärjestyksessä viimeiseksi palkkojen, korkojen ja verojen jälkeen (ts. osakkeenomistajien vaateet). Kysymys yritysten sidosryhmien huomioon ottamisesta ilmeisesti on ratkaistava niin, että sidosryhmien, kuten työntekijöiden, etujen varjelemisesta tulee yritykselle kannattava investointi, ei ulkoa määräytyvä velvoite.

Myös velkojien roolia voidaan korostaa. Koska vararikko aiheuttaa liikkeenjohdolle merkittäviä kustannuksia, yrityksen on tuotettava riittävästi kassavirtaa, jotta velat voidaan hoitaa. Velkarahoitus onkin tehokkaampi kurin luomisessa kuin osakerahoitus.[8]

Omistajavalvontaa, corporate governance -järjestelmää, voidaan myös kehittää. Heikot instituutiot on korvattava paremmilla. Johdon kannustinjärjestelmillä tähdätään parhaan tuloksen aikaansaamiseen. Tuloksesta riippumaton palkkio ei tätä takaa. Optioihin liittyneiden ongelmien vuoksi niitä on aihetta arvioida kriittisesti. Optioetu voidaan pätkiä pitkitetyksi etuudeksi,

joka voidaan lunastaa vain vaiheittain. Vertailuindeksit täytyy ottaa laajemmin käyttöön. Tarkoitus on palkita osaamisesta, ei onnesta. Yhtiöiden hallituksiin voidaan nimittää myös sidosryhmien edustajia.

Myös markkinat voivat ottaa rankaisijan roolin. Kilpailu luo tarpeen rakentaa mainetta (Hörner (2002)) ja hyvä maine hyvästä laadusta taas auttaa ylläpitämään hyvää hintaa (Shapiro 1983). On silti hyödyllistä kysyä, missä lopulta on kuluttajien moraali. Varastettu tavarakin näyttää aina löytävän ostajansa. Kuluttajien halukkuus sitoutua tuotteeseen tavallaan säätelee sitä, mitä arvoja yritykset kokevat itselleen oikeiksi. Globalisoituneessa maailmassa kuluttajat sijaitsevat eri puolilla maailmaa eikä heidän preferenssejään voida koordinoida.

Markkinamekanismi huolehtii ilmeisesti puutteellisesti eettisistä valinnoista. Vapaa markkinatalous ei siis ole arvo sinänsä vaan se vaatii huoltojoukkojen apua. Omistajavallan kehittäminen on tässä käyttökelpoinen strategia.

YRITYSTEN PERUSARVOT

Yritysten perusarvot ovat nousseet pohdittavaksi. On esim. tiedostettu terveen henkilöstöpolitiikan merkitys. Loppuun palaminen ei voi olla elämän tarkoitus. Hyvä työviihtyvyys sen sijaan on arvo sinänsä. Pienipalkkaisen työntekijän motivointi voi olla yhtä tärkeätä kuin johtajan.

Yrityksen eettinen vahvuus näkyy siinä, millä lailla se kohtelee työntekijöitään, asiakkaitaan, rahoittajiaan ja minkä arvoista informaatiota se tarjoaa itsestään omistajilleen, sijoittajille ja verottajalla ja miten se pitää kiinni sopimuksistaan. Organisaation etiikka on ko. organisaatiossa toimivien ihmisten etiikkaa. Imagon rakentamisen tarve on jatkuva.

Monella on se käsitys, että länsimaiden arvojärjestelmässä materialistiset arvot ovat vahvoilla. Kapitalismi näyttää heijastavan näitä arvoja. Väärinkäytöksineen kapitalismi on huono talousjärjestelmä. Suuryhtiöt käyttävät valtaa ja vallankäyttäjiä ei rakasteta. Huononakin kapitalismi on kuitenkin paras tuntemamme järjestelmä. Se on luonut maailmaan uskomattomasti vaurautta. Maailma on riskeistä rakennettu, eivätkä riskit ole kapitalismin syytä. Ihmisen sisäinen onni taas on järjestelmästä riippumaton.

Oikeiden valintojen tekemiseksi ihminen tarvitsee paitsi oikeata tietoa yrityksistä myös riittävää kykyä jäsentää maailmaamme ja ympäröivää todellisuutta. Mutu-tieto ei riitä. Ideologiat ja oppisuunnat ovat kuin huumetta kadoksissa olevalle ihmismielelle. Väärät profeetat saavat seuraajia ja opetuslapsia. Älyllinen ratkaisu yrityksen arvojen tunnistamisessa lähtee siitä, että

yritykselle maine on tärkeä. Jokainen haluaa samaistua tänä päivänä Nokiaan. Kukaan ei halua samaistua Arthur Anderseniin.

Maailmassa on paljon vääryyttä. Ihminen ilmeisesti ei muutu. Mutta siinä määrin kuin ihmisen toimintaa säätelevät instituutiot ja organisaatiot, kannustimet ovat ihmisen luomia, niitä voidaan kehittää ja parantaa ja tätä kautta tehdä maailma paremmaksi paikaksi.

KIRJALLISUUTTA

Akerlof, G. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", The Quarterly Journal of Economics, 84, 3, 488-500.

Aghion, P. and Bolton, P. (1992): "An Incomplete Contracts Approach to Financial Contracting", Review of Economic Studies, 59, 473-494.

Coase, R. (1937): "The Nature of the Firm", Economica 4, 386-405.

Etiikka ja talous (2003), Vesa Kanniainen – Matti Sintonen, toim. WSOY.

Grant, S., King, S. P. and Polar, B. (1996): "Information Externalities, Share-Price Based Incentives and Managerial Behavious", Journal of economic Surveys, 10, 1-21.

Geroski, P.A. (1995): "What Do We Know About Entry", International Journal of Industrial Organization 13, 421-440.

Greenan, N. and Guellec, D. (1994): "Coordination within the Firm and Endogenous Growth", Industrial and Corporate Change, 3, 173-197.

Cremer, J. (1993): "Corporate Culture and Shared Knowledge", Industrial and Corporate Change, 2, 351-386.

Frank, R.H. (1988): Passions within Reasons, New York: W.W.Norton & Co.

Friedman, M. (1970): "The Social Responsibility of Business is To Increase Its Profits", The New York Times, September 13, (Quoted from Thomsen (2001)).

Hall. B.H. (2002): "The Financing of Research and Development", NBER wp. 8773.

Hart, O. (2001): "Financial Contracting", Journal of Economic Literature XXXIX, 1079-1100.

Hermalin, B.E. (1999):"Leading for the Long-Term", working paper, Walter A. Haas School of Business, University of California, Berkeley.

Hermalin, B.E. (2002): Economics & Corporate Culture, University of California, Berkeley, February.

Holmström, B. (1979): "Moral Hazard and Observability", Bell Journal of Economics, 10 ,74-91.

Holmström, B. (1982): "Moral Hazard in Teams", Bell Journal of Economics, 13, 434-340.

Holmström, B.R. and Tirole, J. (1989): "The Theory of the Firm", Handbook of Industrial Organization, 1, eds. R.Schmalensee and R.Willig, Noth-Holland.

Hörner, J. (2002): "Reputation and Competition", American Economic Review, 644-663.

Jensen, M.C. and Meckling, W. (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Capital Structure", Journal of Financial Economics 3, 305-360.

Jensen, M.C. (1986): "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers ", The American Economic Review, 76, 2, 323-329.

Kandel, E. and Lazear, E.P. (1992): "Peer Pressure and Partnerships", Journal of Political Economy, 801-817.

Kreps, D.M. (1990): "Corporate Culture and Economic Theory", in J.E.Alt and K.A.Shepsle, eds., Perspectives on Positive Political Economy, Cambridge, England: Cambridge University Press.

Lazear, E.P. (1995): "Corporate Culture and the Diffusion of Values", in H. Siebert, ed., Trends iin Business Organization, Tubingen, Germany: J.C.B.Mohr (Paul Siebeck).

Lindbeck, A. and Snower, D. (2000): "Multitask Learning and Reorganization of Work: From Tayloristic to Holistic Organization", Journal of Labor Economics, 18, 353-376.

Modigliani, F. and Miller, M. H. (1958): "The Cost of Capital, Corporation Finance and the Theory of Investment", American Economic Review, 48, 261-297.

Myers, S.C. and Majluf, N.S. (1984): "Corporate Financing and Investment Decisions when Firms Have Information that Investors Do Not Have", Journal of Financial Economics, 13, 187-221.

Ridley, M. (1996): "Origin of Virtue", Sane Töregård Agency.

Rob, R. and and Zemsky, P. (2002): "Social Capital, Corporate Culture, and Incentive Intensity", RAND Journal of Economics, 33, 243-257.

Shapiro, C. (1983): "Premiums for High Quality Products as Returns to Reputations", Quarterly Journal of Economics, 659-679.

Scharfstein, D.S. and Stein, J. (1990): "Herd Behavior and Investment", American Economic Review,. 465-79.

Shumpeter, J.A. (1934): "The Theory of Economic Development", Cambridge: Harvard University Press.

Stein, J.C. (1989): "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior", Quarterly Journal of Economics, 104, 655-69.

Stulz, R. (1990): "Managerial Discretion and Optimal Financing Policies", Journal of Financial Economics 26, 3-27.

Tirole, J. (2001): "Corporate Governance", Econometrica, 69, 1-35.

Thomsen, S. (2001): "Business Ethics as Corporate Governance", European Journal of Law and Economics, 11, 153-164.

Williamson, O. (1975): Markets and Hierarchies: Analysis and Antitrust Implications, New York: The Free Press.

Wilson, E. (1975): Sociobiology. The New Synthesis, Massachusetts: Harvard University Press 2000.

Zingales, L. (2000): "In Search for New Foundations", The Journal of Finance, 4, 1623-53.

LOPPUVIITTEET

[1] Suomessa ovat olleet näkyvästi esillä mm. Uotin veljesten yli 7 vuotta kestänyt oikeudenkäynti, virkavaltaa pilkkanaan pitäneen liikemies Fryckmanin pakomatka, jota kansa on saanut seurata kuin jännitysnäytelmän jatkokertomusta, it-sektorin uusyritysten rahastus piensijoittajien kustannuksella juuri sopivasti ennen it-kuplan puhkeamista v. 2000, tulosvaroitusten antaminen pörssisääntöjen vastaisesti liian myöhään, eräät sisäpiirikaupat ja niitä koskevat epäilyt, ahvenanmaalaisen Ålandsbankenin hallintoneuvoston jäsenten merkintäoikeuskaupat, poliisitutkimukset pörssiyhtiö Benefonin tiedottamisesta ja ns. Gibraltar-tapauksen oikeuskäsittely.

[2] Laajaksi osoittautui kapitalismin pettureiden etujoukko USA:ssa, mukana todellisia jalokiviä. Enron, Arthur Andersen, Worldcom, Merck, AOL Time Warner ja monet muut. Myös Xeroxissa paljastui jättimäisiä epäselvyyksiä. Euroopassa vastaavaa ei ollut ilmaantunut ennen kuin hollantilaisen maailman kolmanneksi suurimman vähittäiskaupan yrityksen Aholdin todettiin peukaloineen tulostaan vuosina 2001-2002. Tosin Ruotsissa ABB:n johdon salaiset eläke-edut olivat tätä ennen pysäyttäneet avoimuutta ja tasa-arvoa vaalineen naapurimme kansalaiset.

[3] Näitä kysymyksi käsitellään laajemmin kirjassa Etiikka ja talous, 2003, toim. Kanniainen V. ja Sintonen M.

[4] Kotimaisista tuoreista esimerkeistä käyvät esim. pörssiyhtiöt Comptel ja Orion, joiden riskit ovat realisoituneet. Edellinen oli pari vuotta sitten tähtiyritys. Tuoreen tilinpäätöksen mukaan se tekee suurta tappiota. Orion tekee voittoa, mutta vuosia kestänyt satsaus uuteen ahdistuslääkkeeseen osoittautui toiveajatteluksi.

[5] Itse asiassa yrityksen luonteen ymmärtäminen laajemminkin on ollut haasteellinen tehtävä, ks. Coase (1937), Holmström and Tirole (1989).

[6] Toistuvuuden idea avaa mahdollisuuden ns. toistuvien pelien teorian hyötykäyttöön tälläkin tutkimuslohkolla. Toki on tunnettu tosiasia, että toistuville peleille on tyypillistä tasapainojen monikäsitteisyys. Tämä viittaa siihen, että tarvitaan koordinaatiomekanismi, joka mahdollistaa tasapainon valinnan. Yrityskulttuuri voisi tässä toimia hyödyllisenä koordinaatiomekanismina.

[7] Lauman mukana meneminen voi johtaa mittaviin menetyksiin, mutta sillä on myös oma rationaalinen selityksensä (Scharfstein and Stein 1990).

[8] On siksi kiintoisaa havaita, että esim. Soneran umts-kaupat rahoitettiin ainakin tästä näkökulmasta oikeaoppisesti velalla, pitkän aikavälin tavoitteita painottaen. Soneran kivikkoinen tie ei rakentunut lyhytnäköisyydelle vaan epäonnistumiselle markkinoiden kehityksen arvioimisessa. Soneran umts-hanke oli investointi optioon, joka antoi pääsyn uuteen teknologiaan.

7. The Venture Capitalist as a Competent Outsider

Gunnar Eliasson*

Abstract:

The innovator integrates technologies in new and creative combinations. The entrepreneur identifies the commercially interesting innovations and the venture capitalist understands, recognizes and provides reasonably priced financing. A deep secondary market is needed to allow the venture capitalist to exit profitably and pass on the winning innovations to the industrialist for large-scale production and marketing.

At all stages specialist industrialist competence is needed to achieve a successful end result. And expectations about the end result determine incentives at earlier stages of the innovation and entrepreneurial process. I find that the diversity of specialist competencies, the number of venture capitalists and the organization of the secondary exit market are particularly important and much neglected links in the early industrialization phases of a new industry. The venture capital industry is fairly new and fully developed only in the U.S., notably in the San Francisco Bay area, where it has been, and still is instrumental in the development of IT and biotech industries. High taxes are observed to have reduced the supply of independently wealthy and industrially experienced venture capitalists in Europe that make up the core competence base of the US venture capital industry. The substitute government supply of risk capital has been washed clean of the same competence, and, furthermore, tends to be too slow in providing funds for high-tech firms with a fast turnover of technology.

I find that the venture capitalists not only have to be industrially competent to recognize the entrepreneurs and be many, representing extreme diversity of competencies to avoid missing the winners. What looks like a lack of visible entrepreneurship may in fact be a lack of competent venture capitalists and exit markets.

* Gunnar Eliasson is Professor at the Department of Industrial Economics and Management, The Royal Institute of Technology (KTH). This essay is an update of an earlier intermediate report (Eliasson 1997b) with the same title reporting on 29 interviews with Swedish, European, US and Taiwanese small firms and venture capitalists. Since then 38 more interviews with 27 more firms, and several return visits to firms already interviewed, have been carried out in the context of two separate projects, one on the use of highly educated people in production, and one on the role of corporate headquarters (Eliasson 2001). One particular reason for the update has been to investigate the claims that the supply of venture capital in Sweden has now been improved to the extent that Sweden is no longer disadvantaged in this respect.

I also find that despite claims to the opposite the lack of industrial competence in the financial community in Sweden and continental Europe serving the new and small companies in radically new technologies has not improved much since an earlier study (Eliasson 1997b), and the financial bubble in the IT and biotech markets bears witness to that conclusion. It has rather been the case that more money is now being allocated through the same narrow competence filters as before. This is synonymous to a deterioration in the quality of venture capital supplies, and if there is any place in the economy where economies of scale are not present it is in the venture capital industry.

THE PROBLEM

Sustainable and rapid growth in the Experimentally Organized Economy (EOE, Eliasson 1987,1991a) is achieved through the Schumpeterian Creative Destruction process of Table 7.1. New firms challenge existing firms and force them to improve through reorganization and rationalization or succumb to competition and exit. While innovative entry is the mover of the Schumpeterian creative destruction process, also the exit process is critical to activate the desired sustainable growth at the macro level by releasing resources to these successful and growing firms, the winners[1] (Eliasson-Taymaz 2000).

A key factor behind such successful project filtering is the industrial competence residing in the venture capital industry. The dynamic efficiency of resource allocation is guided by *competence bloc theory* being an integral part of the theory of the EOE (Eliasson-Eliasson 1996, Eliasson 2003b). A competence bloc lists the minimum of tacit functional competencies needed to achieve an efficient allocation of resources. At the top of the causal chain resides the *competent customer*, who determines the orientation of technological development. The competence bloc, therefore, has a demand definition that orients the project selection at later stages. We have the *innovator* who integrates technologies in new and creative combinations. The *entrepreneur* identifies the commercially interesting innovations and the industrially *competent venture capitalist* understands, recognizes and provides reasonably priced financing. A secondary *exit market* is needed to allow the venture capitalist pass on the winners to the *industrialist* for large-scale production and marketing. A critical part of the competence exercised in the competence bloc is to determine the optimal time to pass the relay baton on to the industrialist.

The competence bloc is an organization of competent actors that together coordinate functionally defined competencies over markets or hierarchies for the evaluation and allocation of projects. Dynamic efficiency is achieved (see further below) when each project is subjected to a maximum competent and varied evaluation, minimizing the risk of losing winning projects (Eliasson 2003b).

When being criticized for rejecting promising innovative projects the banker often argues that there is no lack of money, only of good investment projects. The reason, however, as we will demonstrate, is more likely to be that the banker does not understand the investment project , notably projects in radically new technologies, and therefore charges too such for advancing the resources. Under such conditions there will be little entrepreneurial and industrial activity based on the innovation and the risk of losing winners will be high. While it is statically efficient to close down bad projects early, it is dynamically inefficient to terminate winning projects. It is established as a trivial conclusion that the latter is most likely to happen with unique, novel projects in market situations where resource providers lack the competence to understand the project, are few and do not have to compete for projects.

In this study *venture capital has been defined as the supply of finance with such embodied complementary industrial competence.* This definition has also guided the interviews. It should be observed that this is a quite limiting definition that excludes much of the so called risk capital covered in statistical surveys of the venture capital market. This distinction will become important for my conclusions.

We take it for granted that to stay competitive, and ahead of the rapidly learning and industrializing non- industrial world the mature industrial countries have to mobilize their own superior industrial knowledge bases for commercialization and industrialization. In that process the funding of unique, high-tech innovations becomes critical.

This essay, hence, investigates the role of the competent venture capitalist in making high-tech entrepreneurship, new firm establishment, industrialization and economic growth.

The organization of the paper is as follows. We begin by establishing the role of the entrepreneur in economic growth, then proceed to clarify the nature of the entrepreneur, and notably (next) the nature of the high-tech entrepreneur who requires unique competencies on the part of the capitalist to be understood, recognized and funded.

Having done that we go on with the main theme of the paper to clarify the nature of the necessary venture capital competence in principle (theory)

and in practice, as revealed by interviews of suppliers and receivers of venture capital over a 10 year period. We conclude with some observations on policy.

THE ENTREPRENEUR IN ECONOMIC GROWTH

The Schumpeterian entrepreneur sees business opportunities that no one has discovered before. He may be the innovator himself or he may be the person who identifies the commercial potential of someone else's innovation. He establishes a business (a firm) on this innovation.

New firm entry, the *rationalization and reorganization* of existing firms and the exit of inferior firms (Eliasson 1995a; see Table 7.1) define the four fundamental investment mechanisms of the Schumpeterian creative destruction process, exit being a disinvestment.

The reorganization and rationalization of firms may be both innovative and defensive responses to competition. Innovative reorganization often gives fast results. Exit is the Darwinian process of clearing the market of low performers and releasing resources for more competent actors.

Entrepreneurial entry yields significant macroeconomic effects only in the long run, but exercises immediate competitive discipline on incumbent actors. This disciplining role was regarded to be of critical importance for macro economic performance already by Adam Smith (1776) who regarded freedom of entry to be the most relevant indication of competition and of the self-regulating invisible hand of markets[2], but was (Eliasson 1991b) relegated to intellectual insignificance by the post second world war macro and neo walrasian economists who had little understanding of the functioning of competitive markets.

The circumstances under which more or less entrepreneurship occurs are therefore – in this essay – critical for the overall long run performance of

Table 7.1. The Four Investments of Schumpeterian Creative Destruction

1. Innovative entry
2. Reorganization of existing firms
3. Rationalization of existing firms
4. Competitive exit (bankruptcy institution)

Note: The source is Eliasson (1995a).

the macro economy. It is particularly important to clarify the origin of incentives behind high-tech entrepreneurship.

THE NATURE OF ENTREPRENEURSHIP

Entrepreneurship as manifest in the form of new firm entry, or more generally (Jenner 1966) new product entry is usually characterized by a wide spread of ex post performance characteristics. The population of new entrants on the average exhibits equal or slightly lower performance than the average incumbent, but a much wider spread in both directions (Granstrand 1986, Eliasson 1991b). This spells out the great uncertainty, as distinct from risk (Knight 1921), associated with entrepreneurship.

The incentives of the entrepreneur have to be sufficient to overcome this uncertainty. The outcome depends on the ability of the entrepreneur in bringing order to an uncertain business situation, such that it is transformed into one of subjectively calculable risks (Eliasson 1990a).

Entrepreneurship is necessary also in big, existing firms. Their "intrapreneurship" is often organizational. Ponder the competence specification of a business organization in Table 7.2. The upper entrepreneurship (orientation) items in Table 7.2 constitute the ability of top management to reorient the firm in a new direction, or the formation of a new business. This entrepreneurship stage involves creating something new, or to follow Schumpeter (1911), something new that cannot be objectively predicted on the basis of past experience. Here business intuition plays a critical role. Intuition can be defined (Eliasson 1996a, p. 135) as a personal understanding of a new situation that cannot be credibly (objectively, scientifically) communicated to outsiders. To the outsiders, hence, the whole affair is shrouded in uncertainty in Knight's (1921) original meaning, but not to the insider, entrepreneur.

The insider entrepreneur has already formed his or her subjective opinion on what should be done. By converting (to the outsider) incomprehensible and prohibitive uncertainty into low and subjectively computable risks the entrepreneur gets his mind in order to go ahead, provided he can get funded.

Entrepreneurship is said to be more typical of some societies than of other, or typical in the past of some country rather than now (Payne 1974, Wiener 1981). The entrepreneur, in a broader sense, is the individual who conceives a new business opportunity and takes an initiative to exploit it. There does not have to be anything high- tech in entrepreneurship. Most entrepreneurship, in fact, occurs in the form of variations on known industrial

applications and requires few faculties of recognition on the part of the banker. Funding can be provided at reasonable costs through non-competent sources. Our concern, on the other hand, is with odd, unique innovations where little prior experience exists in the entrepreneurial and financial markets. Particular experience, competence and creative faculties are required for identifying and recognizing the commercial potential of such projects. We call these innovations "high- tech", because scientifically based complex innovation is the large, perhaps dominant group of innovations requiring such particular entrepreneurial and venture competencies.

THE ROLE OF THE VENTURE CAPITALIST IN ENTREPRENEURSHIP

A country is often said to lack entrepreneurs. This is a current Swedish worry. This may, however, be a complete misunderstanding. A necessary condition for the existence of entrepreneurship ex post is resource provision at reasonable cost. If resources are not available there may very well be plenty of ex ante innovators and entrepreneurs, but little visible ex post presence. This seems to be the case in the Swedish health care industry (Eliasson 2002).

The first stage resource provider, the venture capitalist, not only has to provide financing. His price for that provision is contingent upon the uncertainty (to him) surrounding the new venture, and his or her ability to understand and recognize the business potential of the new venture. If the venture capitalist does not understand the business venture he or she will charge excessively for funding it, to the extent that there is little left for the innovator/entrepreneur. Hence, there will be little incentive for them to go on. The

Table 7.2. Competence Specification of the Firm

Orientation
1. Sense of direction (business intuition)
2. Risk willing
Selection
3. Efficient identification of mistakes
4. Effective correction of mistakes
Operation
5. Efficient coordination
6. Efficient learning feedback to (1)

Note: The source is Eliasson (1996).

ability to understand and competently evaluate the project proposed to him by the entrepreneur distinguishes the venture capitalist from the banker. The incentives for innovators and entrepreneurs, hence, depend critically on the competence of venture capitalists to understand what they are proposing. If there is only one, or a few, as opposed to many venture capitalists, representing widely diversified , embodied industrial competencies to understand and recognize winning projects the less likely that a winning project will find financing.

This talent of recognition is necessary to this funding in the direction of feeding above average innovators and entrepreneurs into the Schumpeterian creative destruction process of Table 7.1. Without such competence, venture capitalists will only function as risk diversifiers. [3] Ex post records on new firm entry would then suggest such a high risk premium that radical new entry would disappear. Only venture capitalists with the competence to recognize radically new business ventures will survive, and it will be possible to characterize such venture capitalists empirically by studying successful entrepreneurial environments.

Not recognizing the great winners, however, is by far the most costly business mistake, far more costly than keeping the bad projects on a little bit longer, since it deprives the economy of large future profits and resources. In the world of the experimentally organized economy that recognizes business mistakes as an important transactions cost item (Eliasson and Eliasson 2002) the function of the competence bloc is to minimize the economic incidence of business mistakes. One could say that *a complete and horizontally varied competence bloc lowers the premature exit of winners*, i.e. before they have been discovered (see below).

The standard view permeating the proponents of central planning and the centralized innovation systems policies of high-tax European economies has been that central overview helps improve the allocation of resources. Even though this view is gradually dying with its holders because of the disastrous experience of the centrally planned economies, soft versions linger on and the national innovation systems approaches are even on the increase being supported by the bureaucracies in Brussels and several European countries.

By using the theory of the EOE and of competence blocs we can show theoretically that we would not have had to wait for the bad experiences of centralized economic policy regimes to materialize to abandon centralist policies. Under the assumptions of the theory of the EOE tacit competence coupled with gross ignorance dominate market behavior. Attempting to cen-

tralize information for improved overview washes out the mass of varied competence that governs the millions of decisions in the market, and result in a reduced understanding of the dynamics of the economy and a dynamically inefficient allocation of resources, notably the industrial knowledge capital itself (Eliasson 2001b, 2003b).

This paper identifies and characterizes the role of the venture capitalist as the provider of competence embodied finance. It is true as many bankers and industrialists argue that there is no lack of money, only of good investment projects. The problem is that bankers have the money, but not the competence to recognize the good projects. Hence, they experience a lack of good projects.

Take one example from the medical field. Prostate cancer is a well defined and sometimes malignant, but often slow moving, often curable cancer. Most standard cures, however, have significant negative side effects, that affect the quality of post surgery life. There is radical prostatectomy (cutting it out), radiation therapy using different delivery methods and chemotherapy. Some doctors argue, however, that only changes in diet may be effective. New high tech and more precise methods, possibly without negative side effects are, however, being developed and are funded by the US venture capital industry. There is cryosurgery that kills the cells through precise freezing. There are biological early diagnostics and targeting methods (see *The Economist*, March 29th 1997). There are photosensitive chemicals targeting of cancerous cells that kill the cells when subjected to cold laser[4] light, etc. And each particular new technique can be developed in different variations, some of which may work, some not. To fund one particular winning project variation the venture capitalist has to embody deep understanding of the method and – to be successful – a good intuition on which particular method will work. To cover a range of development broad enough to turn up a winner, you need to get many venture capitalists entering the market.

The typical biotech pharmaceutical innovation has a gestation period of many years, often a decade or more, due to the long clinical testing period.

The "venture capitalist" who does not understand the project will apply such high discount rates that even the most promising project falls through. He or she then operates as a banker, and if there is no competence in the venture capital market the banker's attitude will dominate. There will be no new industry creation. This is the most likely situation in mature economies (see below). As one interviewed venture capitalist expressed it: venture capitalism in its modern sense is something fairly recent even in the

U.S. It parallels the development of the "modern" new types of industries, and these only exist in volume in the U.S.

WHAT KIND OF UNCERTAINTY?

Each business venture is a more or less adventurous experiment. It is based on an idea of an individual or a team of individuals. This business experiment is tested in the market. The more radically novel the more adventurous, and the larger the possibility of a major mistake.

But the positive side is: The more radical the larger the potential profit, if you are right, and the more so the fewer your competitors. Since the innovating entrepreneur rarely can fall back on his or her own resources to realize the business experiment, he or she first has to run its proposal through a filter of outsider evaluators. It may be the accountant's eyes of the big corporation, the slow process of a public authority or, if available, the competent venture capitalist's demands on a return.

The more radical the project the more difficult it is to evaluate. And the further off the profits the more unreliable the estimated returns. Therefore, two errors are constantly made;

Type I Error; Bad project are kept alive for two long.
Type II Error; The winners are shut down.

Standard investment calculations can be applied to Type I errors and explicit and well defined responsibilities can be meted out for those who have failed. The more radical and the further off in time the more likely that the project will not be found worthy of further funding. Type II errors are rarely discovered, at best close Type II errors (see Eliasson and Eliasson 1996a). It is difficult to make somebody responsible for Type II errors, even though the implicit costs in the form of lost profits are often enormous.

It is also difficult to compensate those who have prevented Type II errors, because the fair compensation would be enormous. Besides, unassociated people will at the time rush in to share in the praise and confuse the evaluation.[5]

This risk bias is all too natural. All formal tests screen for expected outcomes. Unexpected outcomes will, therefore, not be understood by standard evaluations and, hence, are likely to be thrown out. Hence, the standard evaluation is always biased against novelty and in favor of repetitive similarities that preserve a past that can be understood.

The tacit competence of the entrepreneur to identify the commercial potential of an innovation is matched by the tacit competence of the venture capitalist to recognize and fund the winning entrepreneur.

Running and reorganizing a going concern in repetitive production is necessarily fraught with fundamental uncertainty. Many mistakes are done along the way because the business situation has been misunderstood by those who make the decisions. Identifying commercially viable innovations and discovering entrepreneurs are much more demanding on heterogeneous competencies that have to match. *The venture capital market, hence, allows a decentralization of decision making through the market such that innovative entrepreneurship can exist simultaneously with large scale long-term financing*. It, hence, constitutes an innovative market alternative to the hierarchical organization of the large corporation. This is the case in the pharmaceutical industry where large cash rich companies have lost their capacity to innovate. The small innovating firms, on the other hand, do not have the resources to fund their innovations through clinical testing. An intermediate and competent venture capital market exists to fund the uncertain long innovation phase, and a secondary exit market with many industrial buyers to compete for the good strategic innovations and take them on to industrial scale production and distribution are what is needed to bridge that classical gap and pull new industrial ideas forward. KaroBio of Sweden (see Eliasson and Eliasson 1997) is an intermediary trader in such a decentralized market for distributed innovative and venture capital competence in the biotechnology area.

A MODEL OF VENTURE COMPETITION FOR INNOVATIONS

A venture capitalist assumes a significant part of the business risk of the project he gets financially involved in. If he is not familiar with the business himself and if he cannot obtain (for him) credible advise on its nature, he is likely to apply such high risk premia in his project evaluation as to predate on the innovator, make the project unprofitable or remove himself from the project.

He may spread his risks over may projects and hope that one success will pay for the other mistakes, but this only begs the question. The more projects he spreads his risks over the further he removes himself from the individual projects. The less he understands the individual projects the larger the risk premium he applies. It does not improve the quality of the individual project selection that a separate market for risks can be created. The less the resource providers understand the investment projects they are looking at, the higher the risk premium they apply and the fewer of the projects that will

get funded and the higher the likelihood that odd and unfamiliar projects will be rejected. The ability of the resource provider, hence, is a critical determinant of the (subjective) profitability of the project. The conventional separability assumptions of finance theory, hence, do not apply, and, hence, not the consequent conclusion on risk diversification.

The demands on specific project competence on the part of the resource provider make banks and large investment institutes incompetent as inefficient financiers outside the area of standardized projects described in most literature. The same applies to detached public sector policy programs to stimulate investments. The big, resource-using firms occupy a middle position. They have the resources, and are often very competent in recognizing winners within their narrow range of specialties. Since innovative activity will always occupy a broader range of technology than the existing industry, the financial resources of incumbent firms can never substitute more than partially for venture capitalists. *Any country adopting the policy of running the funding of industrial development through big firms and Government financing will experience* a *long-run concentration of their industry to an increasingly narrow range of traditional and to many well known technologies and an increased competitive exposure.*

The economic analysts carry an analytical tradition that makes it impossible for them to understand matters outside their range of theory. The bankers have the same problem. They apply a standard filter to screen investment projects for known (to them understandable) outcomes, and reject what they don't understand. Now take the opposite position. Suppose one venture capitalist understands and recognizes the business potential of an innovation. If no one else does he has no competition and will be able to pick up the innovation at a bargain price. This is very often the situation when somebody in, or in the neighborhood of an existing firm has come up with an innovative idea. With only one "customer" there won't be much money for the innovator. The market is full of examples of employee innovators who have received a token reward for something the firm has then gone on to make billions on. That one customer understands the innovation, therefore, is almost as bad as if none does. To achieve the double purpose of lowering the risk of rejecting the winner and sufficient compensation (incentives) for the innovator/entrepreneur you need both many venture capitalists and many customers representing a broad range of competence.

The Edgeworth box of classical economic theory might serve as a pedagogical illustration. In the Edgeworth box analysis two traders with wares to sell are confronted. Each has a relative "barter" supply price for his

goods in terms of the wares offered. There is usually an intermediate range within which each party will be better off compared to being on the trade-off line. A game of negotiations determines where a trade will occur. This is the typical situation in the markets for unique innovations, where one innovator/entrepreneur with an odd idea is confronted with one possible resource provider with a strong negotiating hand when it comes to setting the conditions for trade. Ponder the alternative situation with one resource provider who understands the unique innovation quite well and many other resource providers ("banks") who don't understand the innovation at all and (the new situation) with more than one resource provider who understand the innovation very well and know very well how to commercialize the project. In the alternative situation competition will follow and push up the compensation for the innovator/entrepreneur. Incentives for innovation and entrepreneurship will increase and innovative resources will be more efficiently allocated. Economic growth will be higher. Generalizing, the creation of new industries requires many competent venture capitalists for unique innovations to be recognized and efficiently allocated through competition.

RECOGNITION, COMPETITION AND INCENTIVES

Entrepreneurship, to use Schumpeter' s 1911 definition means creating something previously not thought of, that was not possible to derive objectively from past records.[6] The new innovation piloted to the market by the entrepreneur, however, does not necessarily have to be very complex and difficult to understand. Instant recognition is often possible. High tech, if I am allowed to use this term in such a way, comes in at the other end. This time the entrepreneur, as before, has to discover the commercial potential of a complex innovation. But there is no instant recognition of the same commercial potential by the resource providers. Hence, for the creation of genuinely new, high-tech industries, also the resource providers have to excel in competence to understand the new technologies or products and to recognize the winners, and without this recognition there will be no visible high-tech entrepreneurs. Hence, there is no coincidence that the resource providers for high-tech industry – the venture capitalists – began to appear parallel to the large-scale commercialization of new scientific discoveries and military technology in the late fifties. US venture capitalists represent entrepreneurship in itself. Not until the early 1980s, however, did the venture capital industry begin to gain momentum and represent large value. It is also the implicit assumption of this paper that the innovative commercial exploitation of scientific discov-

eries represents a unique competitive advantage of the industrialized economies to stay ahead of the emerging new industrial countries that are rapidly learning the earlier monopoly technologies of the Western world –provided they can organize themselves economically, technologically and socially (in competence blocks, see Eliasson, and Eliasson 1996a, 1997) to make this commercialization possible.

Standard Evaluation Filters Terminate Bad Projects Early but Miss the Winners

New, high-tech industry creation means accessing a much broader variety of business opportunities than before (heterogeneity). Besides the more sophisticated technology needed to create innovations, and the more demanding task for entrepreneurs to discover their commercial potential, the capital providers now have to excel in creative recognition. If they don't they will impose uncertainty and use overly high risk premia in their investment evaluations. Winners will be lost, and there will be little incentives left for the innovator and entrepreneur.

Here we run into the first hurdle. Standard evaluation criteria (filters) may be good at closing down losers, but are bad at identifying winners. The banker's standard evaluation formula incorporates little industrial competence. While resource providers (banks) typically find a large variety of projects using standard evaluation principles on all, competent and creative recognition of winners requires a specialized competence. Hence, the economy needs many venture capitalists with a wide range of specialized competencies and many with the same specialities to make recognition feasible. A necessary condition to ensure effective recognition, hence, is a competitive market of venture capitalists.

There Have to Be Many Venture Capital Providers (Pluralism)

Such venture market *competition* is needed to

1. ensure efficient matching of innovators/entrepreneurs and venture capital specialists such that winners do not get lost (recognition, *pluralism*).
2. enforce high rewards for potential winners to keep incentives for future innovation and entrepreneurship high (*incentives*).
3. supply long-term funding at a reasonable price.

Such competition, however, does not emerge spontaneously. Existing venture capitalists may be satisfied to pick easy money from a small group of devoted innovators. Furthermore, they may not understand that what they charge for their recognizing capacity determines incentives for future innovations. Large firms, together with Government, the only providers of competent venture capital in many European countries, of course prefer to pay little. As long as they are alone, as in much of Europe, the market – as one interviewed European entrepreneur expressed it – is stacked against the innovator/entrepreneur in high tech. Another, a Swede, observed that you "have to watch out. Otherwise, after a decade or so you will be left with 5 percent of equity and nothing to say. For the entrepreneur this means neither money nor fun, and those who now own the business provided little more than money".

The *assessment bias against novelty* cannot be resolved with more information and analysis. Hence, the business of the entrepreneur is unique ideas. It will thus be unusual that the entrepreneur will be more than partially understood even in a rich and diversified venture capital environment. Each entrepreneur needs many competing alternative venture capitalists to choose from, not the other way around which is typical of Europe. *To maximize the probability of being understood, each entrepreneur*, we have just concluded, *has to be exposed to many competing venture capital providers with a varied intellectual and practical background*. In such a venture capital market, rich in heterogeneous and creative competencies the capital provider often takes the initiative. The demand for qualification is therefore on the market, not on the individual venture capitalists. This makes the venture competence issue partly a national policy issue.

With lack of money not being the problem, but lack of competence embodied in the money flow, the problem at hand becomes:

1. to identify the nature and origin of that knowledge and its association with flows of financing,
2. to understand how that knowledge is created,
3. to understand how it can be activated (incentives).

Obviously this is largely an empirical problem.

Table 7.3. Sources of Venture Capital

1. Private family (angel) money	Specialized industrial competence, fast, long-term, limited volume
2. Venture investors	More standardized competence, focused on exit
3. Firms	Narrow, fast, large volume, long-term
4. Government	Slow, bureaucratic, not industrially competent, important last resort when markets fail

CREATIVE RECOGNITION AND THE VIABLE VENTURE MARKET – INTERVIEW RESULTS

Comparing the low and diminishing new establishment in Swedish manufacturing (excluding internal firm new establishment) with fragmented data from other countries the following impressionistic picture emerges. Concentrated industrial production to a small number of firms and a large public sector appear to be associated with low new establishment. A large stock market and (in particular) a large private, non-institutional share appear to be associated with variable and large new entry. High rates of entrepreneurship and new establishment in Finland and the U.S. are compatible with these observations, as is the low level in Sweden.

In order to probe deeper behind those rather superficial observations I conducted a series of 32 interviews with some 29 users and providers of venture capital in the U.S. and in Europe during 1994-1996. An intermediate report on the results is found in Eliasson (1997b). During the period 1997-2002 I extended the interviews to some new firms, notably in Europe, and revisited some of the firms already interviewed. One important reason for this follow up study was the claim that the inferior supply of venture capital competence in Europe (See below) had recently been remedied.

It was always universally recognized in these interviews that, by any measure, the US economy, and notably the Californian and Boston area economies, are far ahead of any other economy in high-tech innovation, entrepreneurship and competent venture capital provision. This is certainly the case if venture capital provision is measured per GNP dollar or per capita (Gilson and Black 1997). In fact, the differences between California and the rest of the US economy were often expressed in the U.S. interviews (also see Saxenian 1994) in words very similar to how the differences between the U.S. and the Swedish economy were characterized in the Swedish interviews. The U.S. is used as a positive reference (bench mark) in all comparisons to follow.

Also, the U.S. has a much larger share of high-tech industry than any other economy, and this is particularly the case when the comparison is restricted to the more narrow high R&D and competence-intensive science-based industries, such as parts of IT and telecommunications, software, biotech and pharmaceuticals.

The emergence of a venture capital industry in the U.S. was associated (with a lag) with the growing commercial exploitation of scientific discoveries and military technology. It began in and around Boston in the 1970s (Route 128) but caught momentum to become an industry in the early 1980s in California (Silicon Valley).

One reason for the increased volume was that access to new sources of finance by way of venture capitalists were opened up. Pension funds were supervised by the Department of Labor which checked individual portfolio items separately. In the 1970s the Department of Labor "discovered" the possibility of portfolio management of pension funds and the pension funds were allowed to take small high-risk positions in ventures. This made an enormous difference to the venture market.

The Nature and Origin of Venture Competence

The nature of the competence required is perhaps best understood from the nature of capital provisions in countries where competent venture capital is missing. In several European countries the only sources of financing available for high-tech ventures are:

1. Government
2. Banks
3. Big firms
4. Family.

A typical source of early venture capital in the U.S., notably California, is private individual wealth, notably wealth acquired through an earlier industrial career.

Family wealth means little in countries that have effectively discouraged individual wealth creation, and converted (through taxation) individual wealth into Government wealth (like Sweden). The latter is synonymous to washing out the competence previously associated with the money flow. Likewise, individual wealth acquired through non-industrial activities, like

successful property speculation (typical of Sweden), is not associated with the competence needed in high- tech entrepreneurship.

California, on the other hand, has thousands of multimillionaires who have became very rich through an early industrial career, being partly compensated through bonuses, stock options or share ownerships. These people typically want to stay in business, but do no longer want to be responsible for day-to-day operations. These people have entered the venture capital industry with their experience (competence) and money for two reasons:

1. They believe they can earn a much higher return on competent venture investments than through other standardized investment alternatives through third parties, because of their unique competence.
2. Through the same competence they can supply reasonably low cost financing for high-tech entrepreneurs and thereby open up funding for such entrepreneurs, sometimes being unique winners.

This is a positive sum game for all, making higher (than in alternative investments) than average returns possible for the venture capitalists.

Together, the rich venture capitalists in California alone represent an enormous range of competence, sufficient - more than in any other place – to capture winning entrepreneurs.

This situation is unique to the United States, notably California and the Boston area. The investors already have a long experience from high-tech industries that can now be generalized and converted into investing in similar industries.

The venture capitalists have experience from industrial risk taking. They have managed successes as well as corrected failed business ventures. They therefore have the positive attitude to business risk taking needed in early ventures; an attitude and experience that is normally lacking in the banking and large-scale investment banking community.

The number of venture capitalists in Silicon Valley is very large, which ensures competition for the good projects. This reduces the risk that winners get lost and raises compensation for innovators and entrepreneurs. In fact competition for good projects in California is so tough that venture capitalists are currently finding that they can often get a higher return from investing abroad.

The venture capital industry in the U.S., so defined, is the result of a time-consuming evolutionary process. It was made possible by some early institutional changes in the early (also) evolution of a science-based industry, including a system of higher education open to commercialization of scientific discoveries. The latter openness in the U.S., apparently, was more associated with the private universities than with the public ones (Eliasson 1996b). Obviously, the prerequisites for the rapid development of a science-based industry, and of a competent venture capital industry are not everywhere present, notably in Europe. Because of the long time needed to develop an indigenous venture capital industry a rapid local development in Europe, besides an open mind to the private capitalist principles that apply, probably requires the import of the services of competent venture capitalists in large volume, and subsequent rapid learning of the trade.

The US venture capital industry, in addition, has developed fairly standard procedures to facilitate and speed up the funding process. A standardized letter of intent, the "term sheet" is said to be important in that respect. If both parties can agree on its standard terms the funding process is fast. It is said to be fair to both parties. It includes, however, clauses that can be invoked that one party, notably the entrepreneur, may not like to see included. Again the implicit terms that would be upheld in court are said to be that such clauses are not invoked arbitrarily. Hence, if a "term sheet" can be agreed on fast, the final contract is in practice also agreed on and action can begin.

The incentives for venture capitalists to appear, furthermore, depend on a number of institutional circumstances. First of all, the early venture capitalist usually enters with the intention to stay long term with the project. The arrangements (see below), furthermore, are almost always such that the venture capitalist is actively involved in the business, supplying complementary competence. However, eventually, the venture capitalist will want out (exit), and not at a bargain price. His return to the investment critically depends on the existence of efficient secondary exit markets. The existence of viable exit markets, hence, also positively affects the terms the venture capitalist is willing to offer the entrepreneur, and this raises the incentives for innovation and entrepreneurship at large. The venture capitalist normally wants to leave, when the new business is ready for large-scale industrialization, or when some outsider offers a very good price for his position.

The existence of this exit market and the return to venture investments, hence, depend on institutions that facilitate ownership transactions. In the case of new, small high tech firms the worst impediment would be unreasonable tax laws that take a large chunk of value at each transaction. In general the US is better than Europe when it comes to supporting a business friendly tax and legal climate. But there are also great differences in Europe. A recent study of the European Venture Capital Association (EVCA) ranks England and Ireland at the top, France in the middle and rigid Germany at the bottom. Sweden comes in as number eight, below the middle (*Business Week*, April 28, pp. 38f; *Dagens Industri*, May 8, 2003). The interviews carried out for this project also rank England far ahead of all Europe when it comes to have developed a sophisticated venture capital industry, but still far behind the US, notably California.

How Is Wealth Activated by the Venture Capitalist (Incentives)?

Both private individual wealth and sophisticated science and innovation exist in many countries, but sophisticated venture capital markets are lacking in most. Which were the unique circumstances in the U.S. that activated both the innovative potential of science and wealth?

Deteriorating returns in alternative ventures is one answer. Another answer is that the competent venture capital industry paralleled the development of high-tech industry. Rich people with related industrial experience and a willingness to take large risks saw opportunities to earn much higher average returns as venture capitalists than as passive investors through third parties.

This positive situation in the U.S. may gradually disappear through increased competition among venture capitalists, lowering the relative availability of opportunities. This currently appears to be the case to some extent, inducing US venture capitalists to look for opportunities abroad.

The positive situation in the U.S., on the other land, cannot easily be recreated elsewhere, say in Sweden. Providing Government funding would not help, if the accompanying competence is missing. And if the critical innovative, science-based industry does not exist, the situation is dead. Lowering the expected returns on other investments increases the relative return to venture investments. In that sense, the high tax, inflationary situation through much of Europe, throughout the 1970s and the 1980s was particularly bad for long-term venture investments, inducing instead a flow of funds into property speculation and similar investments.

Much more than a broad-based, competent project evaluation by the early funder is, however, needed to ensure low-cost venture funding of high tech projects. The start-up venture capitalist usually stays on for some time, but he is rarely willing to stay on forever, to become a passive investor and he may not have the management competence to support later industrialization phases. His competitive advantage, that gives him a high return on his investment once he has identified a winner, is his ability to supply critical complementary management competence support at an early stage, in selecting executive talent etc. He very much acts as an active owner. A critical phase in the value creation chain, however, occurs when the relay race baton is to be passed on to the industrial managers with competence to move the project on to industrial scale production and distribution. It is the task of the venture capitalist to manage that transition through an initial public offering (IPO) or through selling directly to an industrial investor. The existence of a viable, competent secondary market is critical for the venture capitalist to earn a high return and (as well) to pay the innovator/entrepreneur well. His reputation (in the market) – his "brand name" – figures importantly here to signal a low risk to the industrially incompetent participants in an IPO issue.

The venture capitalist wants to exit when the time and the price is right for him or her. Without competent venture capitalists there would not be any entrepreneurs. Without functioning exit markets there would not be any venture capitalists. And without viable exit markets the ability of new, high-tech industries to go industrial scale and/or large firms to make strategic acquisitions to compensate for lacking internal innovative capability would be drastically reduced. Functioning exit markets first developed in the U.S. and exist only in rudimentary form elsewhere. In countries politically keen on reducing and/or controlling private wealth formation, such market activities have been reduced by taxes. Since a viable exit market raises the returns to investment all the way back to the innovator it is of critical importance for the evolution of sophisticated new industries.

In fact, many venture capital firms specialize not only in the early funding and the lining up of investors, but also in organizing the process of going public. Broad, competent and well functioning markets for initial public offerings (IPOs) exists only in the U.S. Such offerings to a large number of anonymous investors, require a reputable issuer. The necessary "brand name" takes a long time to build, and represents large value for the firm.

The critical importance for the entire incentive chain in the creation of new industry appears to be the ability of the markets to select winning projects and to provide a competent and high valuation of the work of innovators and entrepreneurs. Pluralism in venture markets certainly increases the probability of competent evaluation of individual entrepreneurs.

THE US VENTURE CAPITAL MARKETS

The US venture capital market typically has three types of investors;

1. the angels
2. the start-up venture capitalists
3. the mezzotint investors.

Angel money is usually close family or acquaintance money. The start-up investors are the important first professionals who come in at an early stage, when the product is not yet ready but the innovator/entrepreneur needs "bigger money" to go on. The mezzotint investors come in just before the firms are going public. With mezzotint investors the venture capital provisions are becoming more standardized and institutionalized. They provide the "brand name" needed for the entrepreneur/new firm to be credible in the stock market.

The Start-up Venture Capitalist – Evaluation

The start-up venture capitalist is the one who allows for customized attention. He, the promoter, usually takes a 2-5 percent position and invites parties to take limited positions of the same kind to get a more professional evaluation of the project. Additional "big money" from institutions (insurance companies, pension funds and universities etc., not banks) may be invited later.

The start-up investor has one or two specialties related to his background experience; electronics, software, biotech etc. His focus is normally even more narrow, and whenever he spots a project he likes that is "out of focus" he invites partners that know the field better.

This investor, once he has decided that "he likes the project" is normally in for the long term. His evaluation of the project and the terms of agreement he negotiates are based on his understanding of the venture. Projects are often of the kind that a normal banker would not understand and back away from; no pilot product, no sales, only a 5-year history of costs (in

biotech it normally takes 5-10 years to see a profit). Evaluating this situation and the uncertainties involved requires understanding a range of factors (technology, markets, management etc.) and above all experience.

The Nature of Specialist Competence

Understanding high tech and odd projects demands competence. The early venture capitalists normally came with their own industrial experience. They meet groups of sophisticated people who have left the development labs of large high-tech IT firms (like Texas Instruments or HIP) being frustrated by disinterested managers, or young academics in university biotech laboratories, who have developed a new synthesizing method and want to realize a private profit. These innovators, half entrepreneurs lack important experience, notably in management and in identifying economic opportunities. Sometimes they have a good sense of the market potential. The innovations they "offer", however, are often technically or scientifically very difficult. The venture capitalist needs a matching understanding to go on. He or she may like and trust the group, and that may be even more important. To get a better understanding of the project, the venture capitalist usually invites partners.

Complementary Knowhow to Create the Right Multidisciplinary Organization for Business

But evaluation is never only analytical. The start-up venture capitalist always provides complementary knowhow. The innovator/entrepreneur is normally a technical person, sometimes a marketing man, but he or she often lacks management experience. The start-up venture capitalist sits in on the board and is instrumental in supporting the hiring of executive talent for the project. This participation in control is part of the critical uncertainty reduction that turns the project into a reasonable and calculable risk. This active involvement in the business of the entrepreneur is, of course, not frictionless. Two or more self centered individuals are often confronted and part of the venture competence is to manage that personalities' interaction productively. As important is the existence of a deep venture market that offers many choices for the entrepreneur of whom to team up with.

The task of the promoting (leading) venture capitalist is to create the right multidisciplinary organization to make a viable business out of a promising innovative entrepreneurial project and organize its funding.

The venture capitalist also has to excel in a rare type of competence combination. He or she has to be very broad when attempting to select the winners. But once the selection has been made he or she has to focus on the task of going commercial, and not allow other promising projects to detract attention from the chosen business. In addition, once the active owner has turned supporting manager he has to make sure that all necessary competence gets recruited to staff the business organization.

EUROPE VERSUS THE US

Any comparison between US and European venture capital markets have to be at best impressionistic. The interesting differences are not quantifiable, and many more interviews than I have currently done are required for any firm statements. Certain observations, however, kept returning in almost all interviews. There are clear differences in business mentality between Europe and the US. In the US you try and fail, recognizing that in a competitive environment you fail anyhow now and then, and especially if you do nothing.[7] Failure often means stigmatization in Europe, and if so, does not foster a business-minded attitude to risk taking.

Another disturbing feature of Europe is the lack of industrial competence in European venture capital firms. While the US venture capital firms are staffed by rich and industrially experienced people the Swedish venture capital firms appear to be staffed by young business administration graduates that are not independently rich. There was even a lack of engineers or scientists on the staff to assess very difficult technical propositions.

A protective environment, furthermore, makes Europeans less competitive. Social security does not force individuals to act prematurely and take risks to protect themselves economically. Logically, however, it should be the other way around. Social security should cover your business failure economically and make you more willing to take on risks. Apparently this is not the case. This paradox might be explained by the relatively smaller expected returns to engage in risky ventures, compared to the safe incomes or social security cheques that the system offers. This, if so, could then only depend on a very high risk aversion and/or a low industrial high-tech competence. The collapsing welfare economies in Europe then should be a positive sign, since reduced welfare will force reluctant and risk avert people to do something constructive in the market.

Public money, furthermore, dominates European venture capital. The attitude towards such funding was skepticism among US start-up firms.

There was little supplementary competence to be gained from having a Government bureaucracy as a partner. It slowed things down, and above all, when a new innovation reached a rapid development stage, the innovation could more easily be imitated. Then slow Government funds presented an additional risk.

Even more seriously, efficient and deep exit markets do not exist in Europe, a circumstance that lowers expected returns for venture capitalists and entrepreneurs all the way down to innovators. Regulations and tax laws are obvious reasons.

The European industrial scene, furthermore, is more traditional than in the U.S. People build more seniority, move less and tend to get stuck in organizational ranking patterns. Above all, the industrial structure is traditional, large scale mature industries with a conservative knowledge base that is, of course, also mirrored in the venture capital industry. This is, of course, a severe impediment to facilitating the entry of firms based on radically new technologies which requires the ability of all to form and participate in multidisciplinary business organizations that constantly reorganize. In Europe engineers and large-scale repetitive production dominate compared to the U.S. which exhibits a broader and more unstable pattern of functional dominance.

Is the Competence Gap between the US and Europe Narrowing?

The claim in recent years has been that the gap in sophistication and competence between the US and Europe has been narrowing. Before the great IT and biotech bubble that collapsed in 2001 there certainly was a rapid increase in the supply flow of risk finance. One could of course argue that the bubble itself bears witness to the lack of industrial competence in the entire financial community in Europe. The same bubble, however, also occurred in the US, so it is no evidence on this particular problem. As far as Sweden goes, several interviews point to the following. There has been a large increase in the supply of venture funds from the mid 1990s, to and around 2000, that suddenly dried up after the collapse, wrecking several good business propositions (See Eliasson 2001a). But there has been no corresponding increase in the industrial competence base of the venture capital industry. Hence, more money seems to have been allocated through the same narrow competence filters as before, which should rather signal a reduced quality of the venture capital provision, and scale economies is definitely not a typical feature of this industry. Hence the hope that the European venture capital industry that "has

grown up a lot over the past five years" expressed by the current chairman of the EVCA, will have diminished the "long-standing gap in resources and sophistication" between Europe and North America (Business Week, April 28.2003, pp37f) is not supported by the interview evidence of this study.

SUMMARY

A strong theoretical case can be built for the critical importance of a competent venture capital industry to make the creation of new and non-traditional (high tech, science-based) industries possible. Venture capital embodies both technical competence – what is needed to understand and evaluate the new entrepreneurs – and complementary management competence to support the commercialization process.

To avoid losing winners and to provide low-cost equity financing to the entrepreneur, the competence base has to be very broad. To maximize the exposure of each entrepreneur to a competent evaluation many venture capitalists with a varied background are needed.

This is reflected in the composition of the US venture capital industry, and in California in particular; people who have become rich through an industrial career, who now want to get out of operational responsibilities, but still reap a high return on their competence capital through working as active owners. Such a venture capital industry of sufficient magnitude and variation takes a long time to develop.

The big business cash flow is not a satisfactory substitute, since its competence is too narrow, and rather conserves old than builds new structures. Government venture money does not work because it has been cleaned of the competence that was originally embodied in its source, before tax. It is, furthermore, often polluted by irrelevant requirements related to political, regional and employment objectives.

There are some very clear empirical and theoretical results to report on:

1. Venture capital matters tremendously for industrial development and reorganization. It is, however, not the money, but the broad industrial competence associated with the venture money that matters.
2. The entire incentive structure for innovation and entrepreneurship builds on the competence of venture capital providers to select the winners and evaluate innovations and entrepreneurs. Hence, the importance of many

heterogeneous venture capitalists to maximize the exposure of many heterogeneous innovators/entrepreneurs to competent evaluation.

3. The ability of the venture capitalist to pay well "upstream" depends on the existence of credible exit markets where he or she can unload successful ventures at a handsome profit.
4. A venture capital market dominated by public money and large firms is likely to be incompetent by the standards of points 2 and 3.
5. The absence of viable venture capital and exit markets allows large firms to predate on innovators/entrepreneur, thus, reducing incentives for innovation and entrepreneurship and possibly contributing to a less efficient allocation of innovations than potentially possible. This situation can only be remedied by opening up local venture and exit markets to international competition.

The general conclusion is that a rich and competent venture capital market with many actors is a necessary condition for industrial development and renewal. On the comparison between Europe and the U.S., or more distinctly between Sweden and the U.S. we can observe:

First, the US industrial range is very broad, compared to Sweden. There is more of broad-based competition to recognize new ideas than in Sweden.

Second, the important difference between Sweden and the U.S., however, is that the U.S. has a much broader mix of all kinds of venture capital providing institutions, including wealthy people with a broad industrial experience and the mentality of these people which makes them "active prone". Thus, even with the same range of business technology in operation, the probabilities that odd and unrelated innovations will be commercially discovered is larger in the U.S. than in Sweden. This is also the reason why small high-tech firms tend to move their headquarters to the US

Third, Swedish policy ambitions over the last 40 years have reduced and standardized the competence attributes of money, thus breeding relative incompetence. The main generator of venture capital has been taxation and the transformation of private money into public "venture capital" washing out the associated industrial competence in the process. Another factor (the other side of the same thing) is that rich people have been deprived of their capital and/or left the country and that the people who, nevertheless, got rich, got rich on speculation on the collapse of the Swedish welfare economy, notably from inflationary gains on property. These people did not have the competence and experience to be good venture capitalists.

Fourth, Swedish government has been promoting big business in mature industries for the entire postwar period. This gave at least the illusion of more political control, but an increasingly standardized industrial competence base, making the large firms increasingly narrow and incompetent as venture capitalists. As Glete (1996) observes, it might even have locked Sweden into an inferior future industrial technology.

Fifth, and finally, the claim that the gap in sophistication between the US and Europe in competent venture capital provision has diminished is not indicated by the interviews of this study.

REFERENCES

Andersson, Peter and Larsson, Tommy (1998): *Tetra,* Norsteds Förlag.

Anderson, G.M. and Tollison, R.D. (1982): Adam Smith´s Analysis of Joint -Stock Companies, *Journal of Political Economy,* 90, (6) (Dec), 1237ff.

Braunerhjelm, Pontus, ed., (2001): *Huvudkontoren Flyttar Ut,* Stockholm: SNS.

Cantner, Uwe, Hanusch, Horst and Klepper, Steven, eds., (2000): *Economic Evolution, Learning and Complexity,* Heidelberg/New York: Physica Verlag.

Carlsson, Bo, ed., (1997): Technological Systems, Cases, Analyses, Comparisons. Boston and Dordrecht: Kluwer Academic Publishers.

Carlsson, Bo, ed., (2002): *Technological Systems in the Bio Industries- an international study* Boston-Dordrecht-London: Kluwer Academic Publishing.

Eliasson, Gunnar (1977): Competition and Market Processes in a Simulation Model of the Swedish Economty, *American Economic Review,* 67(1: 277-281.

Eliasson, Gunnar, ed., (1978): *A Micro-to-Macro Model of the Swedish Economy,*IUI Conference Reports, 1978:1 Stockholm:IUI.

Eliasson, Gunnar, (1987): Technological Competition and Trade in the Experimentally Organized Economy, IUI Research Reports Nr.32, Stockholm: IUI.

Eliasson, Gunnar (1990a): The Firm as a Competent Team, *Journal of Economic Behavior and Organization,* 13 (3), 175-298.

Eliasson, Gunnar (1991a): Modeling the Experimentally Organized Economy, *Journal of Economic Behavior and Organization,* 16 (1-2), 153-182.

Eliasson, Gunnar (1991b): Deregulation, Innovative Entry and Structural Diversity as a Source of Stable and Rapid Economic Growth, *Journal of Evolutionary Economics,* (1), 49-63.

Eliasson, Gunnar (1995a): The Economics of Technical Change – The macroeconomic consequences of business competence in an experimentally organized economy, *Revue d 'Economie Industrielle,* Numero Exceptionnel.

Eliasson, Gunnar (1995b): *Teknologigenerator eller nationellt prestigeprojekt*? (A Technology Generator or a National Prestige Project? The Swedish Aircraft Industry), Stockholm: City University Press.

Eliasson, Gunnar (1996a): *Firm Objectives, Controls and Organization,* Boston/Dordrecht/ London: Kluwer Academic Publishers.

Eliasson, Gunnar (1996b): The Use of Highly Educated People inProduction; Swedish, German and US experience, KTH, TRITA-IEO 1996:10.

Eliasson, Gunnar (1997a): General Purpose Technologies, Industrial Competence Blocs and Economic Growth; in B. Carlsson (1997).

Eliasson, Gunnar (1997b): The Venture Capitalist as a Competent Insider, KTH TRITA-IEO R 1997:06.

Eliasson, Gunnar (2001a): Huvudkontorets roll i Hälsoindustrin, in Braunerhjelm (2001).

Eliasson, Gunnar (2001b): The Role of Knowledge in Economnic Growth, in Helliwell (2001).

Eliasson, Gunnar (2002): The Health Care Competence Bloc, in Carlsson (2002).

Eliasson, Gunnar (2003a): Global Economic Integration and Regional Allocations of Competence, *Industry and Innovation,* Vol.10, Nr. 1 (March); 75-102.

Eliasson, Gunnar (2003b): Dynamically EfficientAllocations of Knowledge in the Experimentally Organized Economy, Paper presented at the conferenced on *Economic Behavior and Organization* in Honor of Richard H.Day, USC, LA April 26-27,2003.

Eliasson, Gunnar and Eliasson, Åsa (1996a): The Biotechnological Competence Bloc, *Revue* d *'Economie Industrielle,* 78-4~, Trimestre.

Eliasson, Gunnar and Eliasson, Åsa (1997): The Pharmaceutical and Biotechnological Competence Bloc; in B. Carlsson (ed.) (1997).

Eliasson, Gunnar and Eliasson, Åsa (2002): *The Theory of the Firm and the Market for Strategic Acquisitions* Paper presented to the International Joseph A. Schumpeter Society conference in Gainesville Florida, March 28-30,2002.

Eliasson, Gunnar and Taymaz, Erol (2000): Institutions, Entrepreneurship, Economic Flexibility and Growth-experiments on an evolutionary model, in Cantner-Hanusch-Klepper (2000).

Gilson, Ronald J. and Black, Bernard S. (1997): Venture Capital and the Structure of Capital Markets - Banks versus Stock Markets, Mimeo, Stanford University School of Law.

Glete, 1. (1996): "Entrepreneurs and social elites; some reflections on the case of Sweden"; Paper presented at the 6th International Joseph A. Schumpeter Conference, Stockholm June 2-5, 1996; published in Eliasson-Green (eds.), The Micro Foundations of Economic Growth. Ann Arbor: University of Michigan Press, 1988.

Granstrand, Ove (1986): A Note: On Measuring and Modelling Innovative New Entry in Swedish Industry; in R.H. Day and G. Eliasson (eds.), The Dynamics of Market Economies. Stockholm: IUI and Amsterdam: North-Holland.

Helliwell, John, ed., (2001): The Contribution of Human and Social Capital to Sustained Economic Growth and Well- Being, HRDC, Canada and OECD.

Jenner, R.A. (1966): An Information Version of Pure Competition, Economic Journal, 76 (302) (Dec.), 786-805.

Knight, Frank (1921): Risk, Uncertainty and Profit. Boston: Houghton-Mifflin.

Payne, P.L. (1974): British Entrepreneurship in the Nineteenth Century, MacMillan.

Saxenian, Annalee (1994): Regional Advantage; Culture and Competition in Silicon Valley and Route 128, Cambridge MA and London: Harvard University Press.

Schumpeter, Joseph A. (1911, English edition 1934): The Theory of Economic Development, Harvard Economic Studies, Vol. XLVI, Cambridge, MA: Harvard University Press (originally published in German).

Schumpeter, Joseph A. (1942): Capitalism, Socialism and Democracy, New York: Harper & Row.

Wiener, Martin, J. (1981): *English Culture and the Decline of the Industrial Spirit 1850-1980,* Cambridge, London, etc.: Cambridge University Press.

Interviews, in alphabetical order:

Abingworth, London (1997)

Actica Omsorg, Gothenbourg (2002)

Art Wise, Ulm (1995)

BioCare Corporation, Hsinchu Scientific Park, Taiwan (1997)

Brookfield Motor Sports Centre, Cambridge (2000)

Boreas, Burlington, MA (1995)

Camuras, Lund (2001)

Capio, Gothenbourg (2001, 2002)

The Carlson Companies, Minneapolis (1995)

Cellular Data and Multimedia, Lund

CFS Management, Uppsala (1997)

Deutsche Bank Research, Frankfurt a M (1995, 2003)

Diacrin, Charleston, MA (1995)

Discovery Ventures LLC, Sand Hill Road 3000, Menlo Park, CA (1997)

Effnet, Luleå (2000)

Ericsson Mobile Telecommunications, Ideon, Lund (1998)

Eurona Medical, Uppsala (2000 twice, 2001)

Fortune Consulting Group, Taipei (1997)

Gemini Genomics, Cambridge (2000)

The Generics Group, Cambridge (2000)

Genetech Inc., San Francisco (1995)

Glenwood Managerial, 3000 Sand Hill Rd, Menlo Park, CA (1997)

HAl Financing, Los Angeles

Hambrecht & Quist, Taipei (1997)

Hambrecht & Quist, San Francisco (1997)

Hansa Medical, Malmö (2001)

Hasselblad, Gothenbourtg (2000)

High Density Drives, Bandhagen (Stockholm, 2000)

Heraus Holding, Hanau (1995)

Interactiva, Ulm (1995)

IGR, New Scientific Ventures, Cleveland, Ohio (1994)

KaroBio, Huddinge, Sweden (1996 twice, 1999)

Merrill Lynch, London (1997)

Nokia, Helsinki (1994)

Orbital Research, Cleveland, Ohio (1994)

Oxford Technologies, Oxford (2001)

Prelude Trust, Cambridge (2000)

Quality Science, Cleveland, Ohio (1994)

PDT Inc., Santa Barbara, CA (1997)

Perbio Science, Hälsingborg (2000)

Personal Chemistry, Uppsala (2000,twice)

Peregrine Ventures, Los Angeles (1995)

Pyrosequencing, Uppsala (2000 twice, 2001 twice)

Reson, Copenhagen (1999)

Scandinavian Merchant Group, Stockholm (1997)

Score, San Franciso (1998)

Seal Systems, Stockholm (1997)

Signal Pharmaceuticals Inc., San Diego (1995)

Softcare, Ängelholm (2002)

Swedestart, Stockholm (2000)

Therexsys, Staffordshire (1995)

Thinking Machines Corporation, Cambridge MA (1994)

Time Space Radio Systems, Ideon, Lund (1995, 1997, 1998)

Virtual Genetics, Stockholm (2001)

Vocalis, Cambridge (2000)

Zenna&Stärker GmbH&Co (1995)

Together 67 interviews have been carried out with 56 firms.

ENDNOTES

[1] This is the way *endogenous growth* occurs in the Swedish micro-to-macro model MOSES (Eliasson 1977,1978,1991a) which approximates the theory of the EOE. This means that we can support our claims to large potential productivity gains at the macro level through reallocations of resources through the Schumpeterian creative destruction process (Table 7.1) by reference to simulations on that model. See for instance Eliasson and Taymaz (2000).

[2] This efficient self-regulation was disturbed "by monopoly produced by Government regulation" (Anderson and Tollison 1982) which kept low performers alive, locking up resources.

[3] Hence, competence comes embodied with true venture capital. Hence, the venture capitalist cannot unload that uncertainty as a risk to specialist risk managers in the market. If he can he is no longer (by definition) a venture capitalist.

[4] This was the new technology being developed in PDT Inc. (See list of interviewed firms).

[5] There is only one way to capture the innovator's gain, and that is to take on the entrepreneurial role. Compare Lotus (Eliasson 1996a in Chronicle), Ericsson Mobile Telephony (Eliasson 1995b, chapter 11), and Losec, recently the world's most selling prescription drug (Eliasson and Eliasson 1997). All three successes required special circumstances to become a reality. One such special circumstance was one devoted and insistent individual, subsequently worth billions to the shareowners of the firm. The origin of the enormous Rausing family wealth (Tetrapak) is said to be a creative individual in the company, who got very little for his idea. The family has, however, been working hard to create the myth that a family member conceived the innovation (Andersson and Larsson 1998).

[6] This situation often occurs in an economy with chaotic, non-linear properties (Eliasson 1991a).

[7] The interview on p. 34 in *Fortune's* (July 7, 1997) theme article, "Fast money – What makes Silicon Valley go", illustrates the point.

8. LASTENLASTEMME TALOUDELLISET MAHDOLLISUUDET

Jaakko Kiander*

> "...if we consistently act on the optimistic hypothesis, this hypothesis will tend to be realised; whilst by acting on the pessimistic hypothesis we can keep ourselves for ever in the pit of want." John Maynard Keynes: Essays in persuasion (1931)

John Maynard Keynes julkaisi vuonna 1930 kuuluisan esseen *"Economic Possibilities for Our Grandchildren"*. Hän hahmotteli siinä kaukaisempaa tulevaisuutta, jossa varallisuuden karttuminen olisi vähitellen nostanut elintasoa siinä määrin, ettei ihmisten tarvitsisi enää keskittyä pelkästään taloudellisen toimeentulonsa turvaamiseen. Keynesin mielestä taloustieteen tehtävänä oli edistää kasvupolitiikkaa, jonka avulla voitaisiin vähitellen päästä yltäkylläiseen runsauden yhteiskuntaan. Riittävä vauraus vapauttaisi ihmiskunnan liiasta työnteosta samoin kuin taloudellisten ongelmien ikävystyttävästä pohdinnasta. Keynesin arvion mukaan tämä olisi mahdollista kahden sukupolven aikana, kutakuinkin tähän päivään eli 21. vuosisadan alkuun mennessä. Vähitellen toimeentulohuolet jäisivät taka-alalle ja keskeisemmäksi tulisi sen pohtiminen, miten saamme aikamme mielekkäästi kulumaan.

Esseen julkaisuvuosi osui aikaan, jolloin maailmantalous oli ajautunut syviin vaikeuksiin. Pörssikurssien romahdusta, hintojen laskuja ja pankkikriisejä seurannut 1930-luvun alun pula ei ollut kaikkein otollisinta aikaa pohtia tulevaisuuden rajattomia mahdollisuuksia. Pikemminkin tuolloin näytti siltä, että talouksien kasvupotentiaali oli lopussa. Monet pitivät tuolloin laman syynä teknologisen kehityksen aikaansaamaa liikatarjontaa ja toisaalta lamaa edeltänyttä liiallista elintason nousua, joka väistämättä kostautui. Usko markkinatalouteen ja demokratiaan horjui yleisesti. Keynesin tulevaisuusoptimismi ei aikanaan vaikuttanut kovinkaan realistiselta.

Keynesin perusargumentti perustui yksinkertaiseen korkoa korolle -ajatteluun; jos henkeä kohden lasketut reaalitulot ja kansakuntien käytettävissä oleva pääomakanta kasvavat joka vuosi vaikkapa kahdella prosentilla, saavutetaan 70 vuoden aikana nelinkertainen tulotaso (Keynes arvioi, että

* Jaakko Kiander on tutkimusprofessori Valtion taloudellisessa tutkimuskeskuksessa (VATT).

100 vuoden kuluessa tulotaso voisi kasvaa 4-8 -kertaiseksi). Tulojen nelinkertaistuminen tuntuu jo sen verran suurelta parannukselta, että sen turvin voi perustellusti kuvitella mahdollisuuksia riittävän muuhunkin kuin raadantaan ja perustoimeentuloon. Nyt voimme todeta, että Keynesin yltiöoptimismi osoittautui alimitoitetuksi; monissa maissa taloudellinen kasvu oli jopa optimistisia arvioita nopeampaa vuoden 1930 jälkeisenä aikana. Monissa maissa – kuten esimerkiksi Suomessa – saavutettiin noin kolmen prosentin suuruinen keskimääräinen kasvuvauhti, mikä merkitsi tulojen kahdeksankertaistumista 70 vuoden aikana.

Mitä tällä kasvulla saavutettiin? Onko samanlaista parannusta enää odotettavissa seuraavien sukupolvien aikana – vai oliko kyseessä ainutkertainen parannus?

TALOUDEN KASVU ON MULLISTANUT YHTEISKUNNAN

Vaurauden kasvu ei onnistunut vapauttamaan ihmisiä työnteosta ja taloudellisista pakoista Keynesin toivomalla tavalla – ei ainakaan kokonaan. Käytämme edelleen valtaosan ajastamme työntekoon ja murehdimme rahojen niukkuutta. Ensi sijassa tämä johtuu siitä, että tulojen kasvaessa myös elintasolle asetettavat minimivaatimukset ovat kasvaneet. Niinpä sosiaalisesti tyydyttävän elintason saamiseksi on työskenneltävä edelleen paljon. Hyväksyttävänä pidettävä minimielintaso sekä sen ylläpitämisen edellyttämät 'kiinteät kustannukset' (näitä ovat paitsi sähkö, vesi ja keskuslämmitys myös riittävä asumisväljyys sekä televisio ja puhelin) ovat nousseet jatkuvasti. Tämä johtuu siitä, että rikkaus ja köyhyys ovat sosiaalisia konstruktioita, jotka syntyvät vertailusta toisiin ihmisiin. Nykyajan köyhä olisi monessa maassa rikas verrattuna 1930-luvun ihmisiin, mutta tämä ei juurikaan lohduta köyhää, joka vertaa itseään omiin aikalaisiinsa.

Yleinen harrastus elintason kasvattamiseen on myös ollut suurempi kuin uskoisi. Reaalitulojen kahdeksankertaistuminen ei ole muodostanut kuluttajille ongelmaa 70 vuoden aikana; tulot on saatu aina kulutettua. Lisääntyneet tulot on käytetty suurempiin ja lukuisampiin asuntoihin, kesämökkeihin, kodinkoneisiin, sisustuselementteihin, autoihin ja veneisiin sekä tietysti myös runsaampaan ruokaan, juomaan ja vaatteisiin, ja lisäksi huvituksiin, matkoihin ja tekstiviesteihin.

Toisaalta Keynesin visio yltäkylläisyydestä ja vapaa-ajan kasvusta on myös toteutunut. Työviikko ei tosin ole Keynesin esittämän arvauksen mukaan lyhentynyt 15 tuntiin, mutta palkkatyön osuus keskimääräisestä eliniästä on kyllä vähentynyt olennaisesti. Tämä ei johdu vain pidentyneistä vuosi-

lomista vaan ennen kaikkea opiskelu- ja eläkevuosien lisääntymisestä. Työelämään tullaan noin viisi vuotta myöhemmin kuin 1930-luvulla, ja työelämästä siirrytään viettämään varsin pitkää eläkeikää. Yli 15-vuotiaat viettävät elämästään työelämässä noin 2/3. Elinaikainen työn määrä on siten vähentynyt suhteellisesti noin kolmanneksen, ja jos vuosityöajan lyhennykset huomioidaan, jopa puolet.

Lisäksi työelämän sisältö ja rakenteet ovat muuttuneet. Suorittavan ja välttämättömyyshyödykkeitä tuottavan työn osuus on vähentynyt – ellei peräti romahtanut – ja erilaisten luovien, viihdyttävien ja hoitavien töiden osuus on lisääntynyt. Huomattavan suuri osa työntekijöistä toimii nykyään tiedonvälityksen, viihteen, opetuksen, tutkimuksen, urheilun, matkailun sekä hoidon ja hoivan parissa. Vuonna 1930 noin 80 prosenttia suomalaisista sai elantonsa alkutuotannosta ja jalostuksesta, eli perinteisestä kovasta työstä. Nyt näiden elinkeinojen osuus on vain noin 30 prosenttia.

Tapahtunut muutos on ollut suuri. 1900-luvun aikana on koettu modernisaatio, suuri teknologinen muutos, vapaa-ajan lisääntyminen, eliniän pidentyminen ja aineellisen elintason moninkertaistuminen. On syytä kysyä, onko vastaavaan – yhtä nopean taloudellisen kasvun jatkumiseen – mahdollisuuksia seuraavien sukupolvien aikana. Millaista tulevaisuutta rohkenemme odottaa lapsillemme tai näiden lapsille?

JATKUVASTI UHANALAINEN TULEVAISUUS

Perinteiset globaalit uhkakuvat

Lähes jatkuvasta elintason ja elämän laadun parantumisesta huolimatta olemme taipuvaisia ajattelemaan, että tulevaisuus on täynnä uhkia. Tulevaisuus onkin jo pitkään ollut kriisissä. Kuuluisuutta saanut Rooman Klubin raportti *Kasvun rajat* kyseenalaisti koko länsimaisen elämänmuodon jatkumisen mahdollisuuden ilmestyessään vuonna 1972. Pessimismi sai lisää uskottavuutta OPECin masinoimasta ensimmäisestä öljykriisistä vuosina 1973-74. Rooman Klubin mukaan öljy ja muut raaka-aineet olivat loppumaisillaan (tunnettujen öljyvarojen piti loppua vuoden 1985 tienoilla). Resurssien ehtymisen vuoksi ei todennäköisenä enää voitu pitää kasvun jatkumista tulevaisuudessa vaan pikemminkin niukkuuden ja kriisien lisääntymistä. Globaali niukkuus merkitsi myös sitä, ettei kehitysmaiden vaurastumiselle juuri näyttänyt jäävän sijaa.

Energian ja raaka-aineiden loppumisen lisäksi meitä on pitkään uhannut myös ympäristökatastrofi. 1960-luvulla virinneen ekologisen kriisitietoi-

suuden kasvattamina olemme oppineet, että maa, vesi ja ilma pilaantuvat vääjäämättä teollistumisen myötä, ja että ympäristömyrkyt uhkaavat terveyttämme. Rachel Carsonin *Silent Springin* sanoma on iskostunut yleiseen länsimaiseen mielipiteeseen. Kalat kuolevat, lajit kuolevat sukupuuttoon ja teollisuuden päästöt sairastuttavat metsät ja ihmiset.

Kolmas tulevaisuuden uhkakuva on pitkään ollut väestöräjähdys. Hallitsemattoman väestönkasvun on ajateltu malthusilaisen logiikan mukaan johtavan väistämättä ruoan, veden ja muiden resurssien loppumiseen. Väestönkasvu olikin lähes äärimmäistä 1900-luvun aikana, jolloin ihmisten lukumäärä maapallolla lisääntyi vajaasta kahdesta miljardista yli kuuteen miljardiin.

1990-luvulla näiden 1970-lukulaisten globaalien uhkien joukkoon liitettiin vielä pelottava visio kasvihuoneilmiöstä ja asteittaisesta vedenpaisumuksesta, jotka ovat odotettavissa 21. vuosisadan kuluessa teollistuneen maailman liiallisten hiilidioksidipäästöjen vuoksi.

Yhteistä kaikille näille uhkakuville on se, että ne joko tekevät taloudellisen kasvun jatkumisen mahdottomaksi tai ainakin niiden välttämisen katsotaan edellyttävän taloudellisen kasvun rajoittamista. Ne ovat myös saaneet aikaan yleisen epäluulon teknologiaa ja taloudellista edistystä kohtaan. Uhkakuvien ympärille on vähitellen muodostunut laaja intressiryhmien verkosto – 'ympäristölobby' – jonka etujen mukaista on pitää yllä kuvaa jatkuvasti pahemmin kriisiytyvästä maailman tilasta. Tunnetuin ja ehkä edustavin esimerkki tällaisesta lienee vuosittaista – aina uutiskynnyksen ylittävää ja aina synkkiä väitteitä sisältävää – maailman tila -raporttia julkaiseva *Worldwatch*-instituutti.

Ympäristöliike ja globalisaation vastainen liike ovat laajalti omaksuneet globaaliuhkiin liittyviä pessimistisiä näkemyksiä, ja niihin perustuvien uhkakuvien avulla ne harjoittavat voimakasta poliittista painostusta kasvua rajoittavien päätösten aikaansaamiseksi. Globaalit uhkakuvat ovat laajalti omaksuttuja ja uskottavia; koulujärjestelmä, tutkijat ja tiedotusvälineet ylläpitävät ja uusintavat niitä jatkuvasti. Uhkakuvien ainoa ongelma on se, että ne ovat epätosia ja perusteettomia, kuten Björn Lomborg vakuuttavasti osoittaa kirjassaan *The Skeptical Environmentalist.*

Uusiutumattomien raaka-aineiden loppuminen ei vaikuta kovin todennäköiseltä. Useimmat raaka-aineet näyttävät riittävän mainiosti, ja monia niistä voidaan kierrättää. Raaka-aineiden reaalihinnat ovat pääsääntöisesti laskeneet, mikä heijastaa pikemminkin ylitarjontaa kuin uhkaavaa pulaa. Energia ei myöskään ole loppumassa: öljyvarojen ehtyminen on todennäköisesti näköpiirissä seuraavan 100 vuoden aikana mutta vaihtoehtojen kehittä-

minen on jo pitkällä. Sähköntuotannossa aurinko- ja tuulivoiman lisäksi ydinvoima tarjoaa mahdollisuuksia tuottaa lähes rajattomia määriä puhdasta energiaa – jostain syystä tämä vaihtoehto ei kuitenkaan näytä kelpaavan niille, jotka eniten ovat huolissaan ilmaston tilasta.

Ympäristön tilan heikkenemisestä ei ole juuri mitään näyttöä. Pikemminkin vesistöjen ja ilman laatu ovat parantuneet viime vuosikymmenten aikana. Ympäristöongelmat eivät myöskään näytä muodostavan pelättyä uhkaa terveydelle. Väestön terveydentila on päinvastoin useimmissa maissa jatkuvasti parantunut ja odotettavissa oleva elinaika pidentynyt. Edes pelätystä väestöräjähdyksestä ei näyttäisi olevan huolta. Tunnetun väestötieteilijän Wolfgang Lutzin mukaan maailman väestönkasvu näyttää hiipuvan 21. vuosisadan puolivälissä. Väestönkasvu on jo lähestulkoon loppunut Euroopassa ja väestö supistuu monissa maissa. Väestönkasvu ei ole johtanut nälkäongelmiin, vaan maailman elintarviketuotanto on jatkuvasti kasvanut väestöä nopeammin.

Vakavimpana uhkana pidetäänkin yleisesti tällä hetkellä lämpenemisen aikaansaamaa ilmastonmuutosta. Arviot muutoksen suuruudesta vaihtelevat kuitenkin paljon. On mahdollista, ellei peräti todennäköistä, että lämpeneminen jää niin vähäiseksi (1-2 asteeseen seuraavan 100 vuoden kuluessa), ettei merkittäviä vaikutuksia ilmene. Lisäksi hieman pidemmällä aikavälillä (muutaman vuosituhannen päästä) todennäköisesti suurimmaksi uhaksi pohjoisen pallonpuoliskon asukkaille muodostuu lämpenemisen sijasta kylmeneminen ja nykyisen interglasiaarin päättävä seuraava jääkausi.

Taloudellinen pessimismi

Aineksia pessimistiseen tulevaisuudenkuvaan voidaan kuitenkin koota myös taloudellisista tekijöistä, vaikka edellä esitellyt globaalit uhkakuvat väljähtyisivätkin tosiasioiden paineessa. Taloudellinen pessimismi ei ole pelkästään suomalainen ilmiö – se on hyvin suosittua kaikkialla Länsi-Euroopassa. Vaikka voisimme olettaa (tai jopa uskoa!), että energiaa riittää, että emme huku saasteisiin ja että väestöräjähdys ei aiheuta hallitsemattomia kansainvaelluksia, voimme silti suhtautua pessimistisesti omiin taloudellisiin mahdollisuuksiimme.

1990-luvun aikana opimme Suomessa olemaan perinteisten uhkakuvien lisäksi huolestuneita myös puhtaasti taloudellisista uhkatekijöistä. 1990-lukua varjosti lama ja käsitys siitä, ettei meillä ole varaa aiemmin saavutettuun elintasoon tai hyvinvointivaltion etuuksiin. Julkinen talous uhkasi mennä vararikkoon. Monien yllätykseksi lamaa seurasi kuitenkin nopean

kasvun kausi, joka korjasi monia ongelmia. Laman jälkeinen tilanne on siksi hieman helpottunut, mutta uusi uhkakuva on liitetty väestön ikääntymiseen ja siitä mahdollisesti seuraaviin eläke- ja hoivamenopommeihin, jotka edelleen uhkaavat elintasoa (koska veronkorotuksia odotetaan) ja hyvinvointivaltiota (palveluiden ja eläkkeiden maksamiseen ei ole ehkä rahaa). Emme myöskään pidä uskottavana, että työttömät voisivat työllistyä tai että suomalaisten työllisyysaste voisi joskus saavuttaa muiden pohjoismaiden tason.

Pessimistinen diskurssi on itseään ruokkivaa ja jopa itseään toteuttavaa. Se pitää yllä yleistä käsitystä, että julkisten palvelujen ylläpitämiseen ei ole varaa, tuleviin eläkkeisiin ei ole varaa, palkankorotuksiin ei ole varaa, eikä juuri mihinkään muuhunkaan. Muualla Euroopassa samantapaista ankeutta ruokkivat sekä Euroopan unioni vakaussopimuksineen että Euroopan keskuspankki. Budjettivajeita vahtiva EU:n komissio katsoo, että Euroopan maiden on – varsinkin laskusuhdanteessa – leikattava julkisia menoja varjellakseen julkisen talouden tasapainoa ja torjuakseen aina uhkaavan inflaation ja epävakauden. EKP vuorostaan arvioi, ettei Euroopan mailla ole juurikaan kasvupotentiaalia, josta seuraa, että veroja ja korkoja ei kannata alentaa, koska kysynnän lisäys ei saisi aikaan muuta kuin inflaatiota. Niinpä onkin parempi olla tekemättä mitään. Asetelma ei ole kaukana Keynesin[1] ironisoimasta Englannin konservatiivipuolueen pulapolitiikasta (tilanteessa, jossa työttömänä oli 10 prosenttia työvoimasta):

> "Ei pidä yrittää työllistää kaikkia, koska se voi kiihdyttää inflaatiota."
> "Ei pidä investoida, koska emme tiedä, onko se tuottavaa."
> "Ei pidä tehdä mitään, koska jonkin tekeminen merkitsee vain sitä, että jotain muuta ei voi tehdä."
> "Olemme jo tottuneet miljoonatyöttömyyteen kahdeksan vuoden ajan ja kaikki on mennyt hyvin. Miksi ottaa muutoksen tuomia riskejä?"
> "Emme lupaa enempää kuin voimme tehdä. Siispä emme lupaa mitään."

Karrikoiden voisi sanoa, että Euroopan unioni, jonka juhlapuheiden mukaan pitäisi olla "maailman kilpailukykyisin talous vuoteen 2010 mennessä", on pysyvästi skleroottinen talousalue, joka säästää ja valittaa, ja toivoo, että USA:n dynaaminen talous vetäisi maailmantalouden taas kasvuun. Eurooppalainen itseluottamus ei riitä siihen, että Euroopan oma talous – joka on yhtä suuri kuin USA:n talous – voisi yhtä lailla vetää itsensä ja maailmantalouden kasvuun. Tämä edellyttäisi riskinottoa, johon eurooppalaista talouspolitiikkaa ohjailevat byrokratiat sen enempää kuin poliitikot varsinkaan suurissa maissa eivät näytä olevan valmiita. Samaan yhteyteen sopii hyvin eurooppalainen teknologiapessimismi, joka pelkää samanaikaisesti ilmas-

tonmuutosta ja vastustaa ydinvoimaa, ja torjuu hysteerisesti geenimanipuloidut elintarvikkeet – ja samalla uuden lupaavan teknologian.[2]

TALOUDELLISEN KASVUN TEKIJÄT

1900-luvun nopea talouskasvu ja tulotason moninkertaistuminen ei ollut selittämätön ihme. Kasvun tekijät ovat identifioitavissa. Mutta voimmeko tukeutua samoihin tekijöihin edelleen 21. vuosisadalla?

Tärkein kasvutekijä oli ja on edelleen kokonaistuottavuuden kasvu. Sen moottorina toimii taas ensi sijassa teknologinen kehitys. On syytä uskoa, että teknologinen kehitys jatkuu myös tulevaisuudessa, jopa entistä nopeampana. Sekä yritysten että julkisen vallan rahoittama T&K-toiminta on laajempaa ja suunnitelmallisempaa kuin koskaan aikaisemmin. Teknologisen kehityksen kasvukontribuutio ei kuitenkaan välttämättä nouse nykyistä suuremmaksi. Tämä johtuu siitä, että monet suhteellisen helposti saavutettavissa olleet tehokkuushyödyt on Suomessa ja muissa edistyneissä maissa jo hyödynnetty. Toisin kuin 1900-luvulla, Suomi ei voi enää 2000-luvulla käyttää hyväksi jäljessä tulijan etua ja kopioida edistyneempien talouksien toimintamalleja.

Toinen taloudellisen kasvun ja tuottavuuden kasvun tekijä on ollut pääomakannan (K) ja pääomaintensiivisyyden (pääoma työllistä kohden, K/L) kasvu. Tämä kasvu jatkui Suomessa aina 1990-luvun alkuun asti. Pääomaintensiivisyys ja samalla myös työn tuottavuus kasvoivat jatkuvasti, kun pääomakanta kumuloitui korkean investointiasteen (I/Y) vuoksi. Korkeaa investointiastetta ylläpidettiin pitkään pääomamarkkinoiden säätelyn ja toistuvien devalvaatioiden avulla. Vanha järjestelmä tuotti toisaalta nopean kasvun ja toisaalta se johti pääoman tehottomaan käyttöön.

Siirtyminen moderniin markkinatalouteen 1990-luvulla on tehnyt tällaisen vanhakantaisen investointistrategian jatkamisen mahdottomaksi. Yritykset ja sijoittajat korostavat pääoman tuottoa, mikä karsii tehottomia investointihankkeita ja johtaa alempaan investointiasteeseen. Ellei suomalaisten sijoituskohteiden tuotto ole selvästi muita kilpailijamaita korkeampi, ei voida myöskään odottaa, että Suomen talouden pääomakanta kasvaisi muita maita nopeammin. Investointiasteen pysyvä lasku hidastaa tulevaisuudessa odotettavissa olevaa talouskasvua ja tuottavuuden kasvua. Suomeen suuntautuvien investointien määrään voidaan kuitenkin vaikuttaa talouspolitiikalla ja suomalaisten omilla päätöksillä, vaikka yksimielisyyttä parhaista keinoista ei ole vielä saavutettukaan.

Kolmas tärkeä kasvun lähde 1900-luvulla oli jatkuvasti kasvanut työllisyys. Työllisten määrä kasvoi paitsi absoluuttisesti myös suhteessa koko väestön määrään. Kasvun taustalla olivat suhteellisen voimakas väestön kasvu, väestön ikärakenteen edullinen muutos (työikäisen väestön osuus kasvoi) ja työhönosallistumisasteen nousu, kun naiset hakeutuivat työmarkkinoille.

Tässä suhteessa ei Suomessa ole kovin paljon aihetta optimismiin. Virallisen väestöennusteen mukaan väestön kasvu pysähtyy 2020-luvulla, ja työikäisen väestön määrä kääntyy laskuun jo lähivuosina. Naisten työhönosallistumisaste on Suomessa jo erittäin korkea. Miesten työllisyysaste voisi ehkä nousta, mutta se edellyttäisi työttömyyden nujertamista, mikä taas on osoittautunut vaikeaksi. Pitkällä tähtäyksellä sekä väestön että talouden kehitys ovat kuitenkin endogeenisia prosesseja.

VÄESTÖKEHITYKSEN JA TALOUSKASVUN TAKAISINKYTKENNÄT

Suomessa samoin kuin muuallakin Euroopassa on jo sopeuduttu ajatukseen siitä, että väestö vanhenee ja supistuu. Tällainen tulevaisuus ei lupaa kovin hyvää. Kasvu hidastuu ja entistä suurempi osa työntekijöiden ansioista joudutaan kanavoimaan kasvavan eläkeläisväestön tarpeisiin. Vaikka tällainen kehitys vaikuttaakin nykyisten syntyvyyslukujen valossa väistämättömältä, ei se kuitenkaan ole ainoa mahdollisuus.

Merkittävän vastaesimerkin tarjoaa Yhdysvallat, jossa väestön odotetaan jatkavan ripeää kasvuaan. Yhdysvaltojen erilainen väestökehitys ei näytä herättävän sen enempää ihmetystä kuin eurooppalainen väestödefaitismi, vaikka syytä olisi: vuoteen 2050 mennessä Yhdysvaltain väestön ennakoidaan kasvavan nykyisestä 290 miljoonasta jopa 500 miljoonaan. Miten tämä on mahdollista samaan aikaan kun Euroopan ja Japanin väestöjen odotetaan supistuvan?

Erilainen kehitys johtuu Yhdysvaltojen suhteellisen korkeasta syntyvyydestä ja laajasta maahanmuutosta. Amerikkalaisten korkeampi syntyvyys vaikuttaa oudolta, koska sitä ei tueta eurooppalaiseen tapaan vanhempainlomilla, lapsilisillä tai julkisella päivähoidolla. Osaselitystä voi etsiä kulttuurisista eroista – Yhdysvalloissa perinteiset keskiluokkaiset perhearvot lienevät vahvemmassa asemassa kuin postmodernissa Euroopassa. Tärkeämpiä selittäjiä voivat olla kuitenkin hyvä työllisyystilanne, korkea tulotaso ja positiiviset tulevaisuuden odotukset. Samoista syistä Yhdysvallat houkuttelee myös jatkuvasti suuria määriä maahanmuuttajia. Voimakas muuttoliike puolestaan tukee väestön kasvua. Suomessa samoin kuin monessa muussa Eu-

roopan maassa lasten hankkimista taas lykätään epävarmana pidetyn ja heikon työllisyystilanteen vuoksi. Työttömyys jarruttaa myös maahanmuuttoa.

Yhdysvaltain odotettavissa oleva myönteinen taloudellinen ja demografinen kehitys herättää kysymyksen, miksi ei tulevaisuudessa myös Euroopassa ja tai ainakin Suomessa voisi olla nykyistä enemmän lapsia ja maahanmuuttajia. Varsinkin Suomen kaltaisessa maassa positiivisen väestö- ja talouskehityksen edellytysten luulisi olevan hyvät.

Maailmassa on toistaiseksi runsaasti potentiaalisia maahanmuuttajia. Suomella on myös hyvät taloudelliset edellytykset muuttua maahanmuuttajien kannalta houkuttelevaksi maaksi. Suomen saavuttama kansantuotteen korkea taso samoin kuin korkea tuottavuuskin osoittavat, että lähivuosina suomalainen tulotaso voi nousta varsin kilpailukykyiseksi, kunhan työttömyys ensin alenee ja työmarkkinoiden vastaanottokyky paranee riittävästi. Suomi on lisäksi suuri ja lähes autio maa, jossa on runsaasti tilaa miellyttäville asuinympäristöille.

Suomeen ennakoitu aleneva väestökehitys perustuu oletukseen, että hedelmällisyys eli suhteellinen syntyvyys pysyy tulevaisuudessa nykyisellään. Keskimääräinen suomalainen nainen synnyttää keskimäärin noin 1,7 lasta. Luku on eurooppalaisittain kyllä hyvä, mutta väestön uusiutumisen kannalta riittämätön. Hedelmällisyysluvun vähäinenkin nousu saisi aikaan tuntuvan muutoksen pitkän ajan väestökehityksessä. Realistinen tavoite voisi olla hedelmällisyysluvun nostaminen lähelle väestön uusiutumista vastaavaa tasoa.

Työllisyys, maahanmuutto ja syntyvyys riippuvat paitsi toisistaan myös muista tekijöistä; näistä keskeisin on yleinen taloudellinen kehitys. On helppo kuvitella tulevaisuus, missä mitään positiivisia muutoksia ei tapahdu. Voidaan ajatella, että nykyiset ongelmat – korkea työttömyys ja kireä verotus yhdessä hyvinvointivaltion rapautumisen kanssa – johtavat negatiiviseen ja itseään ruokkivaan kierteeseen, jossa työllisyys ei jatkossakaan kohene, mutta väestö ikääntyy ja verotus kiristyy.

Negatiivinen kehä

Odotukset väestön ikääntymisestä saavat aikaan sen, että julkinen talous ja kotitaloudet varautuvat tuleviin meno- ja veropaineisiin lisäämällä säästämistään jo nyt. Tämä vähentää taloudellista aktiviteettia, ylläpitää työttömyyttä ja estää verojen alentamisen. Juuttuminen hitaan kasvun ja korkean verotuksen huonoon tasapainoon voi johtaa siihen, että maahanmuuton sijasta yritykset ja niiden kannalta keskeiset työntekijät alkavat muuttaa maasta

pois. Investoinnit supistuvat, mikä hidastaa myös tuottavuuden ja reaalitulojen kasvua. Veropohjan supistuminen pakottaa kiristämään verotusta edelleen, mutta kun kasvua ei ole, on seurauksena hyvinvointivaltion murentuminen. Heikko kehitys voi johtaa myös syntyvyyden laskuun.

Pitkän ajan tasapainossa verotuksen taso riippuu julkisten menojen suuruudesta ja veropohjasta. Jos julkiset menot määräytyvät myös jatkossa samoin kuin nykyään, voidaan arvioida menojen suuruus jonakin tulevaisuuden ajankohtana.[3] Menot on rahoitettava verottamalla tuotantoa tai kulutusta:

$$tY(L,K) = G$$

missä t on veroaste, Y kokonaistuotanto työvoiman (L) ja pääoman (K) funktiona ilmaistuna ja G julkisten menojen taso. Veroasteeksi saadaan:

$$t = G/Y(L,K).$$

Jos julkiset menot ovat 48 yksikköä ja kokonaistuotannon määrä 100 yksikköä, veroasteeksi saadaan 48 %. Jos tuotanto olisi 20 % suurempi, riittäisi samansuuruisten julkisten menojen rahoittamiseen 40 %:n veroaste.

Taulukko 8.1. Negatiivinen kehä

Väestö ikääntyy ja verotus kiristyy
-> Investoinnit supistuvat
-> Yritykset ja osaajat muuttavat pois
-> Väestö supistuu
-> Infrastruktuuri rapistuu
-> Tuottavuuden ja elintason kasvu hiipuu
-> Työllisyys supistuu
-> Verotus kiristyy entisestään
-> Hyvinvointivaltio murenee

Taulukko 8.2. Talouden kaksi tasapainoa

Julkiset menot	Työllisyys	Pääomakanta	Tuotanto	Veroaste	
G	L	K	Y	t	
48	100	100	100	48%	(i)
48	120	120	120	40%	(ii)

Taloudellisen logiikan mukaan tuotannon taso ei ole eksogeeninen muuttuja. Korkeamman tuotannon saavuttaminen edellyttää työpanoksen ja pääomakannan kasvua. Työpanosta voidaan kasvattaa alentamalla työttömyyttä, kasvattamalla osallistumisastetta ja viime kädessä turvaamalla työikäisen väestön kasvu (tai estämällä sen ennustettu supistuminen). Työhön kohdistuvaa verotusta keventämällä voidaan kasvattaa työllisyyttä (kysyntä- ja tarjontavaikutusten kautta) sekä lisäksi vaikuttaa myös väestönkasvun insentiiveihin. Tulevaisuuden työpanos riippuu siten monesta tekijästä, ei vähimmin kasvuodotuksista (E(Y)) tai ricardolais-barrolaisittain ajatellen myös tulevaisuudessa odotettavasta veroasteesta (E(t)):

$$L = L(t, Y, E(t), E(Y), ...).$$

Vastaavasti pääomakannan kehityksen määrääviin investointeihin vaikuttavat myös monet tekijät, joista tärkeimpiä ovat investointien tuottoon vaikuttavat odotettu tuotannon kehitys ja verotus:

$$K = K(E(t), E(Y), ...).$$

Verotuksen keventäminen kasvattaa työllisyyttä ja investointeja. Odotukset tulevasta kehityksestä voivat olla osin itsensä toteuttavia. Jos odotamme, että tulevaisuuden tuotannon taso on 100 yksikköä ja tasapainoveroaste 48 % (tasapaino (i) taulukossa 8.2), on johdonmukaista pyrkiä nostamaan veroaste tälle tasolle. Odotus kireästä verotuksesta vähentää tulevaisuuden tuotto-odotuksia ja rajoittaa taloudellista aktiviteettia, jolloin odotus toteuttaa itsensä.

Positiivinen kehä

Samaisen logiikan mukaisesti voidaan ajatella positiivista tulevaisuudenkuvaa, jossa talouden ja väestökehityksen dynamiikka tukee toinen toistaan. Ajatellaan lähtökohtaista asetelmaa, jossa verotus kevenee ja työllisyys paranee. Muutos kiihdyttää talouskasvua ja parantaa investointikannustimia. Tämän seurauksena investoinnit kasvavat, ja maa (esim. Suomi) alkaa houkutella yrityksiä ja maahanmuuttajia. Muutos johtaa siihen, että vastoin ennusteita väestö jatkaakin kasvuaan. Hyvä työllisyystilanne ja positiiviset tulo-odotukset lisäävät myös syntyvyyttä, kun perheet pitävät tulevaisuuttaan parempana ja turvatumpana. Investointien ansiosta tuottavuuden ja elintason kasvu jatkuu nopeana ja myös infrastruktuuria kehitetään. Myönteisen väestö- ja työllisyyskehityksen ansiosta veropohjat kasvavat voimakkaasti, ja

verotus pysyy keveänä samalla kun hyvinvointipalveluja parannetaan (tasapaino (ii) taulukossa 8.2).

Maahanmuutto edellyttää kilpailukykyistä palkkatasoa ja kasvavaa taloutta, jossa työvoimalle on kysyntää. Sama pätee myös investointeihin. Näitä tavoitteita voidaan edistää yleisen yritys- ja kasvumyönteisen talouspolitiikan ohella myös keveällä verotuksella. Jos talouden näkymät ovat hyvät ja osa verokevennyksistä suunnataan lapsiperheille, ei ole mahdotonta sekään, että syntyvyys nousisi nykyistä korkeammalle tasolle. Investoinnit ja väestönkasvu puolestaan lisäävät talouden kasvupotentiaalia tulevaisuudessa.

Talouden erilaisia ja silti itseään ylläpitäviä tasapainotiloja ('negatiivisia ja positiivisia kehiä') voidaan tarkastella myös taloustieteelle tyypilliseen tapaan funktioita kuvaavien käyrien avulla. Ajatellaan, että verokertymä (tY) riippuu työllisyysasteesta (L), kun veroaste on annettu. Kuvioon 8.1 on piirretty kaksi eri verokertymäkäyrää. Ylemmässä veroaste on korkeampi kuin alemmassa.

Kuviossa 8.1 verokertymä kasvaa kun työllisyysaste nousee. Verokertymää voidaan myös kasvattaa nostamalla veroastetta; siksi t2Y(L) > t1Y(L) kaikilla L:n arvoilla. Tasapainossa julkisten menojen on vastattava tuloja. Niinpä julkisten menojen suuruus riippuu verokertymästä. Kuvioon 8.2 on piirretty kaksi vaihtoehtoista menokäyrää. Ne ovat laskevia; toisin sanoen työllisyyden nousun oletetaan alentavan julkisia menoja, koska tulonsiirtojen tarve työtä tekemättömälle väestölle vähenee.

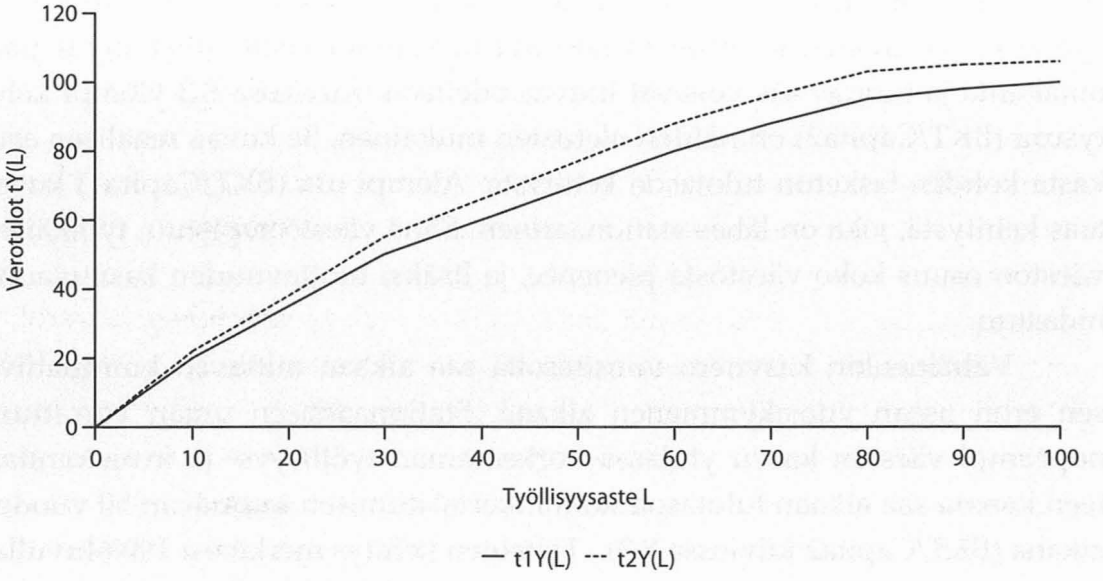
Eri käyrät kohtaavat kuviossa 8.2. Käyrien yhtymäkohdat kuvaavat tasapainotiloja. Ensimmäinen tasapainotila muodostuu pisteessä, jossa työllisyysaste on 60. Korkein tasapainopiste vastaa 80 prosentin työllisyysastetta. Kuviosta voidaan havaita, että alemman työllisyysasteen tasapainossa vero-

Taulukko 8.3. Positiivinen kehä

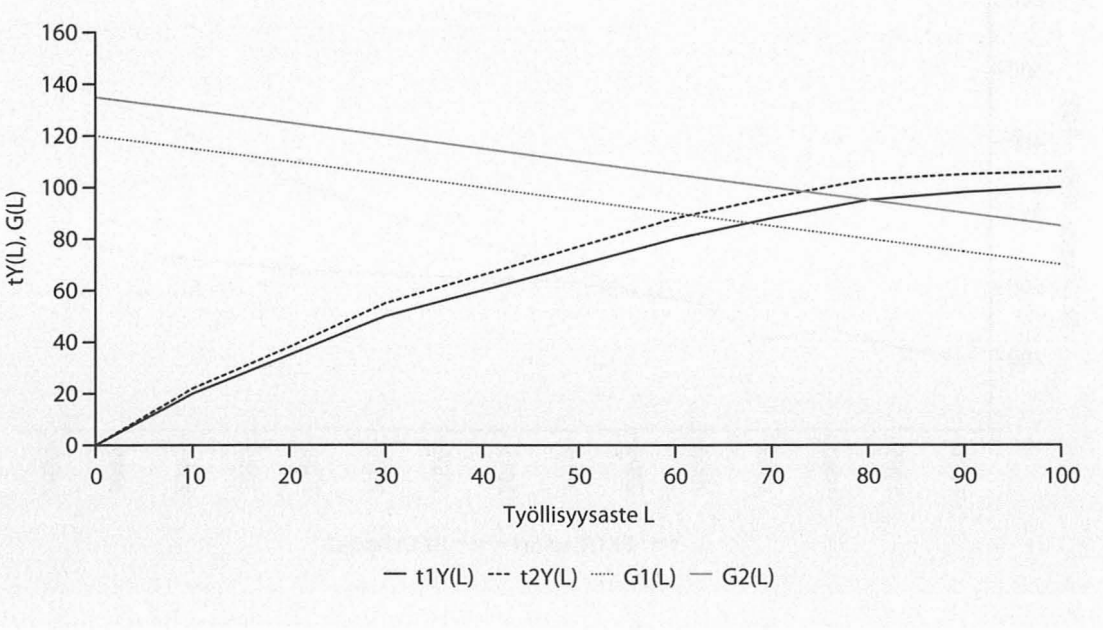
Verotus kevenee ja työllisyys paranee
-> Investoinnit kasvavat
-> Suomi houkuttelee yrityksiä ja maahanmuuttajia
-> Väestö jatkaa kasvuaan
-> Infrastruktuuria kehitetään
-> Tuottavuuden ja elintason kasvu jatkuu nopeana
-> Työllisyys kasvaa
-> Verotus pysyy keveänä
-> Hyvinvointipalveluja parannetaan

aste on korkea (t2) ja julkisten menojen taso matala (G1). Korkean työllisyysasteen tasapainossa veroaste on matala (t1) ja julkisten menojen taso korkea (G2). Työllisyysaste on siten avain sekä kohtuulliseen veroasteeseen että riittävään julkisen sektorin rahoitukseen. Taloudellisen logiikan mukaan on helppo ajatella, että korkean työllisyysasteen ja kohtuullisen veroasteen tasapaino on sopusoinnussa myös nopean tuotannon kasvuvauhdin kanssa.

Kuvio 8.1. Verotulot työllisyys- ja veroasteiden funktiona



Kuvio 8.2. Verotulot ja julkiset menot työllisyys- ja veroasteiden funktiona



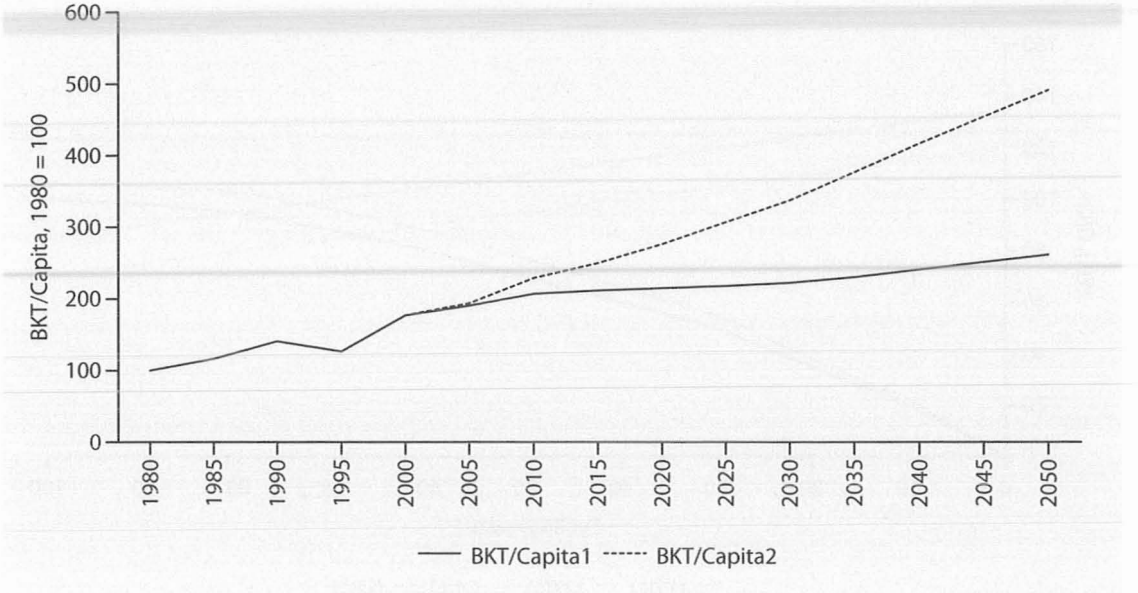
Millä edellytyksillä Suomen talous voisi siirtyä tällaiseen suotuisaan tasapainoon? Kuinka suuria muutoksia nopean kasvun jatkuminen Suomessa edellyttäisi? Tarkastelemme tätä kysymystä seuraavassa väestöskenaarion avulla.

JATKUVA KASVU VERSUS STATIONAARINEN TILA

1900-luvun väestö- ja talouskehityksen jatkuminen Suomessa myös 21. vuosisadalla edellyttäisi edelleen kasvavaa väestöä (maahanmuuton ja kohoavan syntyvyyden avulla) ja riittävän korkeaa investointiastetta, jotta myös pääomakanta ja tuottavuus voisivat kasvaa edelleen. Kuviossa 8.3 ylempi kehitysura (BKT/Capita2) on näiden oletusten mukainen. Se kuvaa reaalisen asukasta kohden lasketun tulotason kehitystä. Alempi ura (BKT/Capita1) kuvaa taas kehitystä, joka on lähes stationaarinen. Siinä väestö supistuu, työikäisen väestön osuus koko väestöstä pienenee, ja lisäksi tuottavuuden kasvuvauhti hidastuu.

Vähäinenkin kasvuero vuositasolla saa aikaan mittavan kumulatiivisen eron usean vuosikymmenen aikana. Stationaariseen uraan verrattuna nopeampi väestön kasvu yhdessä korkeamman työllisyys- ja investointiasteen kanssa saa aikaan tulotason kolminkertaistumisen seuraavan 50 vuoden aikana (BKT/Capita2 kuviossa 8.3). Tällainen kehitys merkitsisi 1900-luvulla

Kuvio 8.3. Reaalinen tulotaso nopean ja hitaan kasvun vaihtoehdoissa



vallinneen kasvuvauhdin jatkumista suurin piirtein entisellään. Sen vaihtoehto on supistuvaan väestökehitykseen ja nykyisen työllisyysasteen jatkumiseen perustuva kehitysura, jossa henkeä kohti laskettujen tulojen kasvu hiipuu vuoden 2010 jälkeen, kun työllisten väestöosuus alenee (BKT/Capita1 kuviossa 8.3). Heikko talouskehitys yhdessä väestön ikääntymisen kanssa johtaa verotuksen kiristymiseen, minkä seurauksena kotitalouksien verojen jälkeen käytettävissä olevat tulot eivät nouse juuri lainkaan.

Stationaarisen uran mukainen hidas kehitys merkitsisi 1900-luvulla vallinneen nopean kasvuvauhdin hiipumista lähelle nollakasvua. Hitaan kasvun vaihtoehdon taustaoletukset eivät kuitenkaan ole epärealistisia. Niissä oletetaan hedelmällisyyden pysyvän ennallaan (syntyvyyden lasku johtuu siten vain ikäluokkien pienenemisestä) ja työllisyysasteen pysyvän vuosien 2000-2003 tasolla. Oletukset nettomaahanmuutosta ja tuottavuuden kasvusta ovat toki pessimistisiä, mutta niitä voi pitää loogisina; huono työllisyystilanne ja heikko tulokehitys eivät houkuttele maahanmuuttajia (vaan voivat johtaa maastamuuttoon), eivätkä heikot kasvunäkymät houkuttele investointeja. Vastaavasti nopean kasvun vaihtoehdossa syntyvyys kohoaa, runsaat investoinnit turvaavat tuottavuuden kasvun, työllisyysaste nousee ja paraneva tulo- ja työllisyyskehitys houkuttelevat runsaasti maahanmuuttajia.

Vaihtoehtoiset oletukset tuottavat perin erinäköiset bruttokansantuotteen volyymin kehitysurat (ks. kuvio 8.4 sivulla 159). Stationaarisen tulokehityksen ura (BKT1) johtaa lähestulkoon bkt:n nollakasvuun vuoden 2010 jälkeen. Nopean kasvun urassa (BKT2) sen sijaan tuotanto kasvaisi 21. vuosisadalla lähes samaa vauhtia kuin edelliselläkin vuosisadalla, minkä seurauksena BKT:n volyymi kolminkertaistuisi vuosina 2000-2050. Tällöin voisi jatkua entisenlainen kehitys, jossa reaalitulot kasvavat selvästi joka vuosi.

Tuottavuuden ohella hitaan ja nopean kasvun skenaariot eroavat väestökehityksensä suhteen. Maahanmuutto ja korkeampi syntyvyys takaavat nopean kasvun vaihtoehdossa sen, että Suomen väestö jatkaisi kasvuaan 2000-luvulla samaan tahtiin kuin 1900-luvulla. Stationaarisen kehityksen

Taulukko 8.4. Kehitysvaihtoehdot vuosille 2005-2050

	BKT1 (hidas)	BKT2 (nopea)
Hedelmällisyysluku	1.7	2.0
Nettomaahanmuutto/v.	0	15000
Työllisyysaste	67%	75%
Tuottavuuden kasvu %/v.	1.1	1.75
BKT:n kasvu %/v.	0.5	3.0

vaihtoehdossa väestö alkaisi taas supistua vuoden 2015 tienoilla, eikä vähenemiselle näkyisi loppua. Vuoteen 2050 mennessä eri vaihtoehtojen erot olisivat kasvaneet jo todella suuriksi, noin miljoonaan henkeen (ks. kuvio 8.5). Puolet tästä erotuksesta johtuu syntyvyydestä ja puolet nettomaahanmuutosta. Molemmat vaihtoehdot ovat mahdollisia; ne mahtuvat stokastisen väestöennusteen vaihteluväliin.[4] Todellisuudessa tuleva väestökehitys ei sen enempää kuin taloudenkaan kehitys ole satunnaista vaikka sen ennustaminen onkin mahdotonta. On varsin todennäköistä, että mahdollinen väestön kasvu tai supistuminen kytkeytyy tietynlaiseen väestökehityksen kanssa sopusoinnussa olevaan taloudelliseen kehitykseen, johon voidaan taas vaikuttaa tietoisilla tulevaisuuteen vaikuttavilla talouspoliittisilla päätöksillä.

LOPUKSI

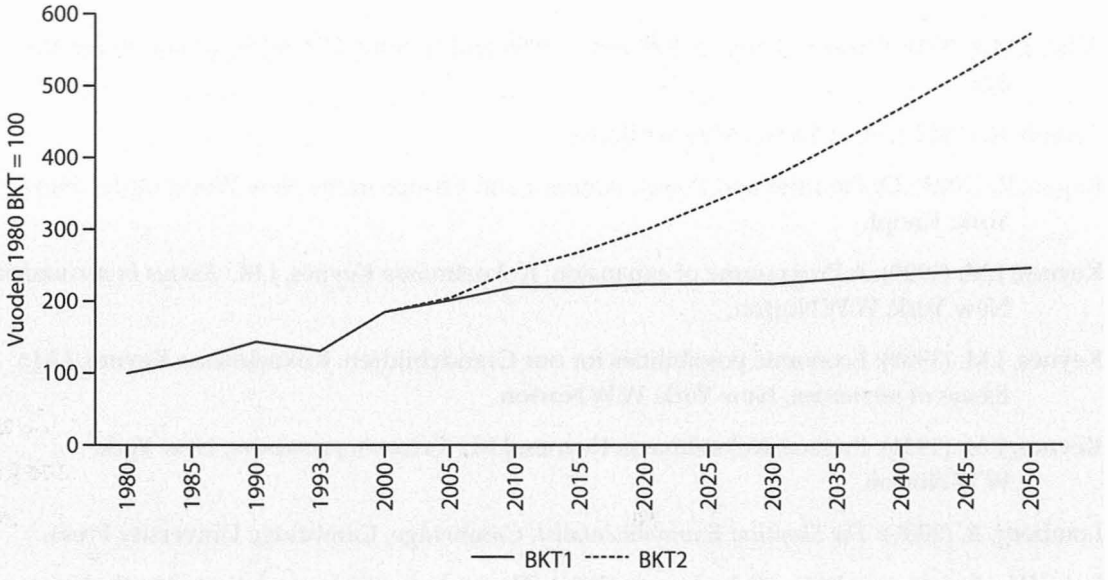
Tässä artikkelissa on pohdittu Suomen kansantalouden kasvumahdollisuuksia tulevaisuudessa. 1900-luvun aikana totuimme nopeaan tuotannon ja elintason kasvuun, joka oli sekä tuottavuuden että työllisyyden kasvun aikaansaamaa. Samanlaisen kehityksen jatkuminen tulevaisuudessa ei ole itsestään selvää. Nykyisten väestöennusteiden mukaan työpanos tulee supistumaan, kun työikäisen väestön määrä alkaa vähentyä. Yhdessä väestön ikääntymisen luomien menopaineiden kanssa tämä voi johtaa kiristyvään verotukseen ja heikentyvään kilpailukykyyn, jotka edelleen heikentävät investointikannustimia ja jarruttavat talouden kasvuvauhtia. Ei ole mahdotonta, että Suomi yhdessä muiden ikääntyvien eurooppalaisten maiden kanssa ajautuu stationaariseen tilaan, jossa talous ei kasva juuri lainkaan. Nollakasvu saattaisi tosin olla monien talouskasvun kriitikoiden mielestä hyvä saavutus.

Mikään kehitys ei kuitenkaan ole väistämätön. Edellä on yritetty hahmotella niitä edellytyksiä, jotka takaisivat 1900-luvulla käynnistyneen nopean talouskasvun ja elintason kohoamisen jatkumisen myös 2000-luvun puolella. Välttämätön edellytys nopealle kasvulle on ennustettua suotuisampi väestökehitys. Tämä on mahdollista, jos syntyvyys nousee hieman nykyisestä ja jos Suomeen suuntautuva maahanmuutto kasvaa. Kumpikaan ehto ei ole periaatteessa mahdoton. Ne edellyttäisivät kuitenkin sitä, että työvoiman kysynnän pitäisi kasvaa selvästi nykyiseltä tasoltaan. Lisäksi ne edellyttävät sitä, että Suomesta tulisi entistä houkuttelevampi toimintaympäristö yrityksille ja maahanmuuttajille.

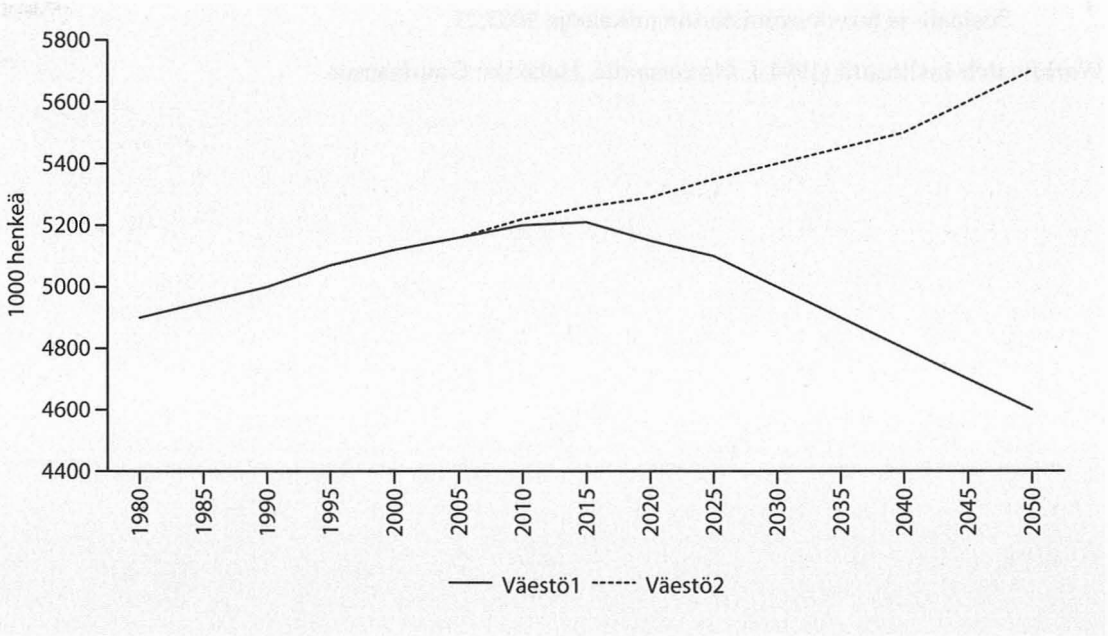
Suomella on hyvät mahdollisuudet yltää muuta Eurooppaa parempaan väestö- ja talouskehitykseen. Meillä on lähtökohtaisesti hyvä kilpailu-

kyky ja vahva julkinen talous. Ne tarjoavat mahdollisuuden kasvua ja maahanmuuttoa suosivaan talouspolitiikkaan.

Kuvio 8.4. Bruttokansantuote nopean ja hitaan kasvun vaihtoehdoissa



Kuvio 8.5. Väestö nopean ja hitaan kasvun vaihtoehdoissa



KIRJALLISUUTTA

Alho, J.M. (2002): The population of Finland in 2050 and beyond. ETLA Discussion Paper No. 826.

Carson, R. (1962): *Silent Spring*. Mariner Books.

Kagan, R. (2003): Of Paradise and Power. America and Europe in the New World Order. New York: Knopf.

Keynes, J.M. (1929): A Programme of expansion. Kokoelmassa Keynes, J.M.: *Essays in persuasion*, New York: W.W.Norton.

Keynes, J.M. (1930): Economic possibilities for our Grandchildren. Kokoelmassa Keynes, J.M.: *Essays in persuasion*, New York: W.W.Norton.

Keynes, J.M. (1931): Preface. Kokoelmassa Keynes, J.M.: *Essays in persuasion*, New York: W.W.Norton.

Lomborg, B. (2001): *The Skeptical Environmentalist*. Cambridge: Cambridge University Press.

Lutz, W. – Sanderson, W.C. – Scherbov, S. (2001): The end of world population growth. *Nature*, 412.

Meadows, D.L. et al. (1972): *The Limits to Growth*. New York: Universe Books.

STM (2002): Sosiaalimenojen kehitys pitkällä aikavälillä. SOMERA-toimikunnan taustaraportti. Sosiaali- ja terveysministeriön julkaisuja 2002:21.

Worldwatch-instituutti (1994-): *Maailmantila*. Helsinki: Gaudeamus.

LOPPUVIITTEET

[1] Keynes: A programme of expansion (1929).

[2] Länsi-Euroopalle tyypilliseksi tullutta passiivis-pessimististä asenneilmastoa ruotii mm. Robert Kagan (2003).

[3] Ks. esim. SOMERA-toimikunnan mietintö, STM (2002).

[4] Ks. tarkemmin Alho (2002).

9. Olemmeko maailman parhaimmistoa? Kansallisen kilpailukykymme puntarointia

Matti Pohjola*

Meitä suomalaisia pidetään vaatimattomina ja vähään tyytyväisinä ihmisinä. Tämä ei kuitenkaan päde poliittisiin johtajiimme, jotka ovat aina asettaneet kunnianhimoisia tavoitteita kansakunnallemme. Maailman huipulle pyrkimisestä onkin tullut monen kansallisen instituutiomme toimintaa ohjaava tekijä.

Esimerkiksi Sitra on asettanut toimintansa tavoitteeksi, että Suomi olisi vuonna 2010 kolmen menestyneimmän kansakunnan joukossa. Tämän kirjoitukseni ensimmäisenä päämääränä on selvittää, millä mittareilla ja miten tämän tavoitteen toteutumista tulisi arvioida kansantalouden osalta. Tehtävä vaatii kansainvälisten vertailujen tekemistä. Toisena päämääränä on selvittää, millä keinoin asetettuun tavoitteeseen tulisi pyrkiä, jotta se toteutuisi – ellei nyt vuoteen 2010 mennessä – niin ainakin mahdollisimman pian. Analyysi perustuu tieteellisen tutkimuksen näkemyksiin niistä tekijöistä, jotka vaikuttavat kansantalouksien menestymiseen aineellisella elintasolla mitaten. Tarkastelun ulkopuolelle jäävät näin ollen muut ihmisten hyvinvointiin ja onnellisuuteen vaikuttavat tekijät.

KANSALLINEN KILPAILUKYKY

Miten Suomi pärjää pitkän ajan tarkastelussa? Kysymykseen vastaaminen edellyttää ensimmäiseksi menestyksen mittaamista. Mikä on käyttökelpoinen kansakunnan kilpailukyvyn mittari?

Yksittäisestä yrityksestä on verrattain helppo sanoa, milloin se on kilpailukykyinen. Yritys menestyy silloin, kun se valmistaa korkealaatuisia tuotteita alhaisin kustannuksin. Tämä käsite voidaan yleistää toimialoja koskevaksi, ja kilpailukykyä voidaan mitata markkina- ja vientiosuuksilla. Ongelmallista on käsitteen laajentaminen koko kansantaloutta koskevaksi.

* Matti Pohjola on professori Helsingin kauppakorkeakoulussa.

Suomalaiset yritykset eivät voi menestyä kaikilla toimialoilla näin määritellyllä tavalla. Huonosti hoidettu yritys menee lopulta konkurssiin ja poistuu markkinoilta. Kansantalous ei voi mennä konkurssiin, eikä se voi poistua kansainvälisestä työnjaosta. Sen heikon menestyksen kantavat kansalaiset alentuneena elintasona, minkä 1990-luvun lama meille suomalaisille opetti. Tässä mielessä on väärin samaistaa yksittäisen yrityksen ja kansantalouden edut.

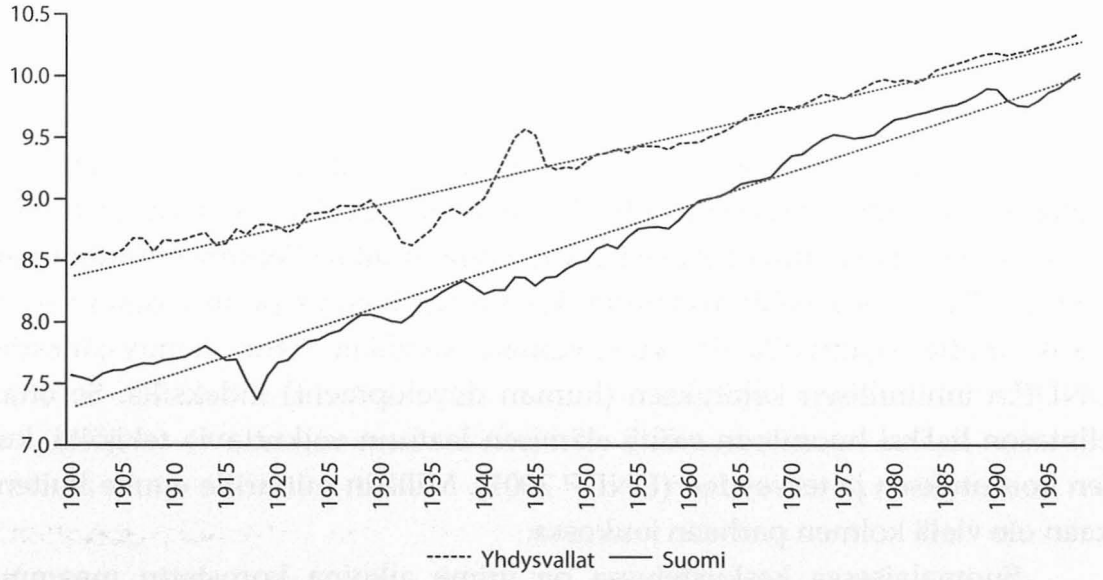
Talouspoliittisessa keskustelussa käytetään bruttokansantuotteen kasvua usein kansantalouden menestymisen mittarina. Näin arvioiden Suomi pärjäsi erinomaisesti 1990-luvun loppuvuosina mutta surkeasti sen alkuvuosina. Tämän mittarin ongelma on kuitenkin se, että kansainvälisissä vertailuissa talouksien kasvuvauhtien on havaittu korreloivan heikosti yli ajan: kansantalous, joka kasvoi voimakkaasti 1980-luvulla, kasvoi eräiden laskelmien mukaan nopeasti viime vuosikymmenellä vain 10-40 prosentin todennäköisyydellä. Japani on tästä hyvä esimerkki, sillä nykyään tämä entinen talouskasvun mallimaa ei enää esikuvaksi kelpaa. Lyhyen ajan kasvuvauhti näyttää määräytyvän kansantalouden ulkoisten tekijöiden vaikutuksesta eli sattuman perusteella.

Maiden elintasoerot ovat puolestaan paljon pysyvämpiä: eri maita kuvaavassa poikkileikkauksessa vuosien 1960 ja 1990 välinen korrelaatio on 90 prosentin luokkaa. Maa, joka oli rikas neljä vuosikymmentä sitten, on 90 prosentin todennäköisyydellä rikas nykyäänkin. Kehittyneiden maiden järjestyskorrelaatio sadan vuoden pituisen periodin yli on peräti 80 prosenttia.

Näistä syistä talouskasvua on tarkasteltava riittävän pitkällä, 20-30 vuoden pituisella ajanjaksolla, jos sitä käytetään kilpailukyvyn mittarina. Kuvio 9.1 näyttää miten asukasta kohti laskettu kansantuote on kehittynyt Suomessa ja Yhdysvalloissa viime vuosisadalla. Se vahvistaa mielikuvan siitä, että olemme saavuttaneet elintasossa Yhdysvaltoja. Kun meillä asukasta kohti laskettu kansantuote kasvoi keskimäärin 2,8 prosenttia vuodessa, niin Yhdysvalloissa kasvu oli 1,9 prosenttia. Kuvion tarkempi tutkiskelu paljastaa kuitenkin, että elintasoeroa kurottiin umpeen erityisen nopeasti 1940-luvun lopulta aina 1970-luvun puoliväliin saakka, mutta sen jälkeen ero ei ole kaventunut, vaan on pysynyt noin 70 prosentissa.

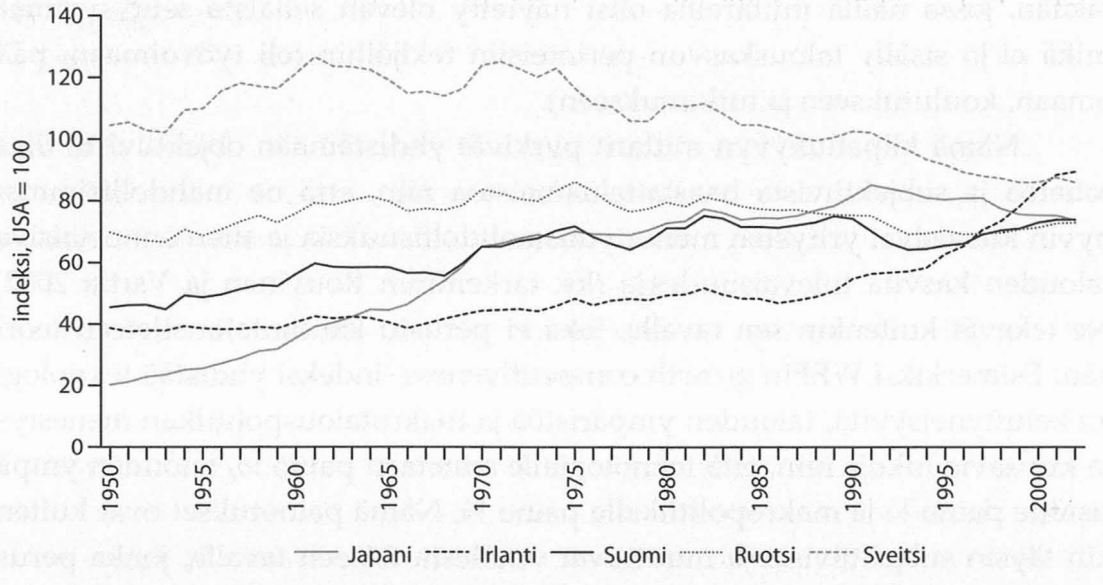
Kuvio 9.2 todistaa tämän asian selkeämmin. Se esittää elintasoa viidessä maassa Yhdysvaltoihin verrattuna vuodesta 1950 vuoteen 2002. Merkille pantavaa on, että elintasoerojen kaventuminen näyttää pysähtyneen Ruotsin ja Suomen osalta jo 1970-luvun lopulla ja Japanin osalta 1990-luvulla. Irlanti – pikemminkin kuin Suomi – on viime vuosikymmenen menestystarina, kun

Kuvio 9.1. Kansantuote asukasta kohti 1900-1998



Kuviossa on logaritminen kansantuote asukasta kohti vuoden 1996 kansainvälisissä dollareissa ostovoimapariteetein arvioituna. Lähteinä ovat Maddison (1995): vuodet 1900-1950 ja Penn World Table 6: vuodet 1950-1998.

Kuvio 9.2. Elintaso Yhdysvaltoihin verrattuna vuosina 1950-2002



Kuviossa on kansantuote asukasta kohti vuoden 1999 kansainvälisissä dollareissa ostovoimapariteetein arvioituna, Yhdysvallat = 100. Lähteinä ovat University of Groningen and The Conference Board, GGDC Total Economy Database, 2003, http://www.eco.rug.nl/ggdc.

menestystä mitataan talouskasvulla. Japanin kokemukset kuitenkin muistuttavat näin lyhyeen ajanjaksoon perustuvan arvioinnin vaaroista. Sveitsi on puolestaan esimerkki maasta, jonka suhteellinen elintaso on laskenut jo pitemmän aikaa. Rikkaus ei siis välttämättä ole pysyvää.

Kansainvälisessä vertailussa Suomen voi katsoa menestyneen hyvin, vaikka elintasoero ei näytä enää kaventuvan Yhdysvaltoihin verrattuna. Olemme kuitenkin saavuttaneet teollisuusmaiden keskimääräisen tason elintason osalta ja jopa ylittäneet sen hyvinvoinnin osalta. Olemme nimittäin sijaluvulla 15 asukasta kohti arvioituun kansantuotteeseen perustuvassa vertailussa, mutta sijaluvulla 10 kun vertailu tehdään YK:n kehitysjärjestön UNDP:n inhimillisen kehityksen (human development) indeksillä. Se ottaa elintason lisäksi huomioon eräitä elämisen laatuun vaikuttavia tekijöitä, kuten koulutuksen ja terveyden (UNDP 2001). Millään mittarille emme kuitenkaan ole vielä kolmen parhaan joukossa.

Suomalaisessa keskustelussa on viime aikoina korostettu maamme menestymistä kansainvälisten laitosten tekemissä kilpailukyvyn arvioinneissa. Sekä International Institute for Management Development -laitoksen (IMD) että World Economic Forumin (WEF) mukaan Suomi sijoittuu toiseksi Yhdysvaltojen jälkeen. Näitä analyysejä ei voi kuitenkaan ottaa kovin vakavasti. Kymmenen viime vuoden aikana on nimittäin julkaistu satoja – ellei peräti jo tuhansia – tieteellisiä tutkimuksia talouskasvusta, mutta en tiedä yhtäkään, jossa näillä mittareilla olisi näytetty olevan sellaista selitysvoimaa, mikä ei jo sisälly talouskasvun perinteisiin tekijöihin (eli työvoimaan, pääomaan, koulutukseen ja tutkimukseen).

Nämä kilpailukyvyn mittarit pyrkivät yhdistämään objektiivista tilastotietoa ja subjektiivista haastatteluaineistoa niin, että ne mahdollisimman hyvin kuvaisivat yritysten menestymismahdollisuuksia ja siten ennustaisivat talouden kasvua tulevaisuudessa (ks. tarkemmin Rouvinen ja Vartia 2002). Ne tekevät kuitenkin sen tavalla, joka ei perustu kansantaloustieteen teoriaan. Esimerkiksi WEFin growth competitiveness -indeksi yhdistää teknologian kehittyneisyyttä, talouden ympäristöä ja makrotalouspolitiikan menestystä kuvaavia lukuja niin, että teknologialle annetaan paino ½, talouden ympäristölle paino ¼ ja makropolitiikalle paino ¼. Nämä painotukset ovat kuitenkin täysin subjektiivisia ja muuttuvat vuodesta toiseen tavalla, jonka perusteita on mahdoton ulkopuolisen ymmärtää.

Suomessa näiden vertailujen sanomaa on tulkittu esimerkiksi siten, että julkinen sektori on edistänyt kansantaloutemme kilpailukykyä. Tämä päätelmä on kuitenkin virheellinen, sillä esimerkiksi WEFin indeksissä oletetaan, että suuri julkinen sektori on aina haitaksi kilpailukyvylle. Suomi sijoittuukin

tämän mittarin osalta vasta sijaluvulle 56 Guatemalan viedessä voiton. Huono sijoitus ei ole kuitenkaan Suomen kokonaismenestyksen kannalta ongelma siksi, että julkisen sektorin paino on vain 6 prosenttia kokonaisindeksissä. Oikea tulkinta onkin siksi se, että kansantaloutemme on kilpailukykyinen suuresta julkisesta sektorista huolimatta. Mikään kansantaloustieteen teoria ei kuitenkaan anna tukea sille väitteelle, että julkisen sektorin koko olisi aina kansantalouden menestymisen este tai sille, että 6 prosenttia olisi sen oikea painokerroin kilpailukykyä arvioitaessa.

Erikoista IMD:n ja WEFin arvioissa on lisäksi se, että Suomi näyttää pärjäävän järjestelmällisesti paremmin subjektiivisilla kuin objektiivisilla mittareilla. Suomalaiset haastateltavat uskovat näin ollen enemmän Suomen kilpailukykyyn kuin amerikkalaiset haastateltavat Yhdysvaltojen kykyyn. Näille arvioille ei voi kuitenkaan antaa paljon arvoa, koska esimerkiksi WEF haastattelee vain noin 20 henkilöä Suomessa (Rouvinen 2001).

Näistä ja monista muista vastaavista tekijöistä seuraa, ettei kilpailukyvyn indekseillä ole kyetty ennustamaan kansantalouksien menestystä. Kuvio 9.3 sivulla 168 näyttää, että talouskasvu oli vuosina 1993-1999 itse asiassa keskimääristä hitaampaa niissä maissa, joiden kilpailukyvyn IMD arvioi parhaaksi vuonna 1993. Hyvänä esimerkkinä on jälleen Japani, jota 1990-luvun alussa yleisesti pidettiin maailman kilpailukykyisimpänä maana. Suomen talous sen sijaan kasvoi hyvin, vaikkei IMD sen kykyyn uskonut. Näiden mittareiden ongelmana onkin yksinkertaisesti se, että ne näyttävät paremmin heijastavan jo toteutunutta kehitystä kuin ennustavan tulevaa.

Miten kilpailukykyä tulisi sitten mitata? Kansantaloustieteen opetus on hyvin yksinkertainen: tuottavuus on ainoa järkevä mittari. Michael Porter (1990) tiivistää asian ytimekkäästi:

> "The only meaningful concept of competitiveness at the national level is *productivity*. The principal goal of a nation is to produce a high and rising standard of living for its citizens. The ability to do so depends on the productivity with which a nation's labor and capital are employed."

Myös Paul Krugman (1994) on samoilla linjoilla:

> "Productivity isn't everything, but in the long run it is almost everything."

Tuottavuutta voidaan mitata kolmella tavalla. Työn tuottavuus kertoo, kuinka paljon tuotantoa saadaan aikaan käytettävissä olevalla työvoimalla. Se on myös kansainvälisesti käyttökelpoinen tulo- ja palkkatason mittari. Talous, joka ei pysty kasvattamaan työn tuottavuuttaan samaa vauhtia kuin

muut maat, köyhtyy väistämättä toisiin maihin verrattuna. Korkean tuottavuuden kansantalous pystyy myös tuottamaan hyödykkeitä alhaisilla kustannuksilla ja silti maksamaan korkeita palkkoja. Tässä mielessä käsite on yhdenmukainen yritystason kilpailukyvyn mittareiden kanssa.

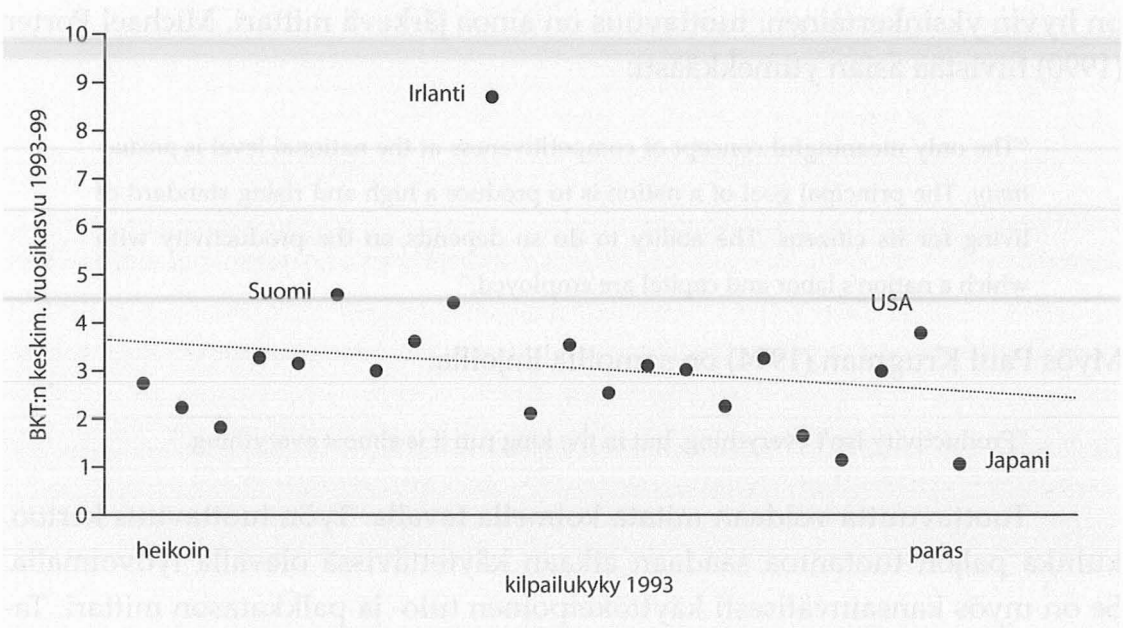
Pääoman tuottavuus mittaa vastaavasti, miten paljon pystytään tuottamaan käytettävissä olevalla pääomakannalla. Näiden kahden osittaistuottavuuden toisistaan erottaminen on tärkeää, koska korkeaa työn tuottavuutta voidaan tavoitella mekanisaatiolla, korvaamalla työtä pääomalla, jolloin pääoman tuottavuus heikkenee.

Kokonaistuottavuus mittaa tuotantoa tietyllä tavalla yhteenlaskettuja panoksia kohti. Korkean työn ja kokonaistuottavuuden kansantalouden menestyminen kansainvälisessä työnjaossa näkyykin korkeana palkka- ja elintasona.

ELINTASOEROJA SELITTÄVÄT TEKIJÄT

Kansantaloustieteen perinteisen uusklassinen kasvuteorian mukaan elintason (BKT asukasta kohti) tai työn tuottavuuden (BKT työpanosta tai työtuntia kohti) kasvuvauhti on pitkällä aikavälillä vakio, joka määräytyy teknologisen kehityksen perusteella. Siihen vaikuttavia tekijöitä ei tämä teoria selitä. Niin sanottu uusi eli endogeenisen kasvun teoria laajentaa perinteisiä oppeja

Kuvio 9.3. Kilpailukyky ja talouskasvu OECD-maissa 1990-luvulla



Lähteinä ovat kilpailukyvyn osalta IMD ja BKT:n kasvu osalta World Development Indicators.

selittämällä teknologian kehityksen tavoitteellisen tutkimus- ja kehitystoiminnan tuloksena. Tämä uusi teoria soveltuu kuitenkin vain kovin pienen maajoukon talouskasvun analyysiin, sillä G7-maat tekevät peräti 90 prosenttia maailman kaikesta tutkimus- ja kehitystoiminnasta. Valtaosaa maailman kansantalouksista voidaan siksi analysoida perinteisellä teorialla olettaen, että teknologia (eli tieto, ideat ja osaaminen) on kehitetty muualla ja että se on siksi näiden maiden kannalta eksogeeninen. Niiden ongelmana onkin, miten hyötyä tästä teknologiasta.

Uusklassinen kasvuteoria soveltuu hyvin elintasoerojen selittämiseen. Vaikka pitkällä aikavälillä kasvuvauhti on vakio, niin rikkaaksi asukasta kohti lasketulla kansantuotteella mitaten kehittyy lopulta kansantalous, joka investoi paljon

1. kiinteään pääomaan eli koneisiin, laitteisiin ja rakenteisiin,
2. henkiseen pääomaan eli koulutukseen,
3. tutkimus- ja kehityspääomaan, organisatoriseen tai sosiaaliseen pääomaan eli niihin kansallisen pääoman komponentteihin, jotka eivät sisälly kiinteään ja henkiseen pääomaan,

tai kansantalous, joka pystyy tehokkaasti

4. omaksumaan uuden teknologian ja siten nostamaan kokonaistuottavuuttaan.

Tämän asian havainnollistamiseksi tarvitaan hieman analyysiä. Tarkastellaan ensin kiinteän pääoman merkitystä elintasoerojen selityksenä. Oletetaan että kansantalouden tuotantomahdollisuuksia voidaan kuvata yksinkertaisella Cobb-Douglas-funktiolla

$$Y_t = AK_t^b (X_t L_t)^{1-b}, \tag{1}$$

jossa Y on kansantuote, K pääomakanta, L väkiluku, X teknologisen kehityksen mittari, b pääoman osuus kansantulosta ja jossa indeksi t mittaa aikaa. Parametrin A ajatellaan kuvaavan kansantalouden kokonaistuottavuuden tai tehokkuuden tasoa. Oletetaan, että teknologia X_t kasvaa vakiovauhdilla g. Soveltamalla tätä tuotantofunktiota perinteiseen Solow'n kasvumalliin saamme tulokseksi, että pitkän ajan tasapainossa elintaso eli kansantuote asukasta kohti Y/L kehittyy seuraavan yhtälön mukaisesti:

$$\frac{Y_t}{L_t} = A^{1/(1-b)} e^{gt} \left(\frac{s}{g+n+d} \right)^{b/(1-b)} \tag{2}$$

jossa n on väestön kasvuvauhti ja d poistojen osuus pääomakannasta ja jossa on oletettu että $X_0 = 1$.

Solow'n kasvumallissa oletetaan, että pääomakantaa lisäävät investoinnit I_t rahoitetaan säästämisellä, joka on vakio-osuus kansantulosta: $I_t = sY_t$. Pääomakannan kasvua kuvaava yhtälö on näin ollen $K_{t+1}-K_t = sY_t - dK_t$. Yhtälö (2) saadaan sijoittamalla tähän kasvuyhtälöön tuotantofunktio (1) ja etsimällä se tasapainoratkaisu, jossa pääomakanta tehokasta työpanosta kohti (K/XL) on vakio.

Pitkän ajan tasapainossa elintaso Y/L kasvaa siis yhtälön (2) mukaan teknologisen kehityksen määrittelemällä vauhdilla g, joka on vakio ja määräytyy mallin ulkopuolelta. Elintaso on puolestaan sitä suurempi mitä korkeampi on säästämis- eli investointiaste s tai mitä suurempi on kokonaistuottavuuden tai tehokkuuden taso A.

Aivan vastaava harjoitelma voidaan tehdä henkisen pääoman, T&K-pääoman, organisatorisen ja sosiaalisen pääoman osalta erikseen tai kaikille yhdessä. Nyt vain s kuvaa investointia henkiseen pääomaan, T&K-pääomaan tai organisatoriseen ja sosiaaliseen pääomaan. Lopputuloksena saadaan edellä kohdissa 1-4 esitetyt väittämät. Se, mikä näistä tekijöistä on paras elintasoerojen selittäjä, selviää vain empiirisessä tarkastelussa.

KASVUTEKIJÖIDEN SELITYSKYKY

1800-luvun alussa rikkaissa maissa asuvat ihmiset olivat keskimäärin vain kaksi kertaa rikkaampia kuin köyhissä maissa asuvat. Nyt he ovat noin 20-30 kertaa rikkaampia. Talouskasvu onkin polarisoitunut: rikkaat maat ovat kasvaneet nopeasti, köyhät hyvin hitaasti, jos ollenkaan.

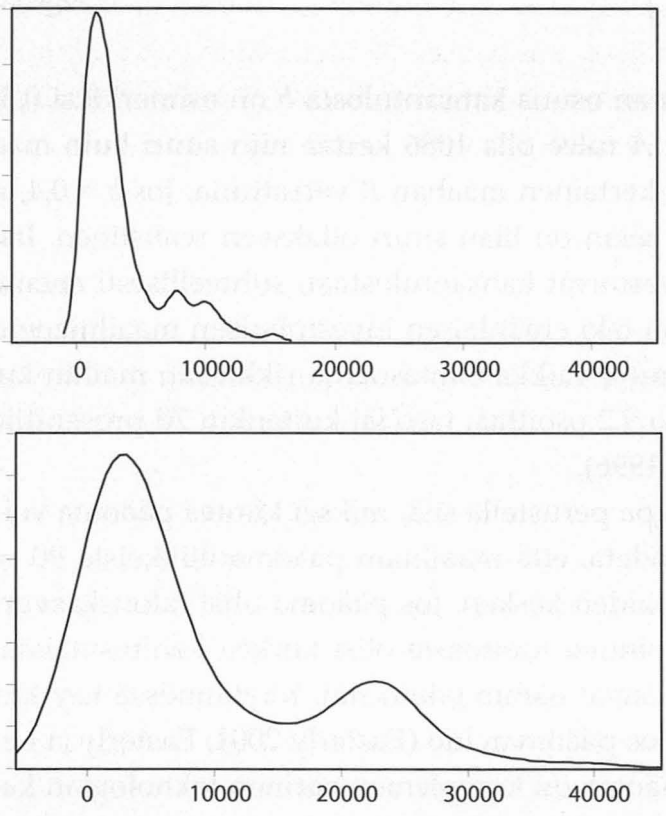
Tämä käy hyvin ilmi kuviosta 9.4, johon on estimoitu asukasta kohti lasketun kansantuotteen jakaumat 108 maan kesken vuosina 1960 ja 1998. Mielenkiintoista on, että jakauma näyttää kehittyvän kaksihuippuiseksi. On syntynyt rikkaiden maiden klubi, jossa elintaso on 20-30 kertaa suurempi kuin muissa maissa keskimäärin. Suomi sijoittuu molemmissa jakaumissa rikkaiden maiden keskikastiin, mikä vahvistaa edellä esitetyn päätelmän. Maamme on onnistunut kasvupolitiikassaan niin, että olemme päässeet rikkaiden klubiin, mutta siinä joukossa emme ole kuitenkaan keskimääräistä parempia.

Mikä edellä mainituista tekijöistä sitten parhaiten selittää näin suuret elintasoerot ja sen, että rikkaat maat näyttävät entisestään rikastuvan köyhien pysyessä köyhinä. Mitä rikkailla mailla on, jota köyhillä ei ole? Vastaus on

kyky omaksua teknologiaa, jota uusklassisessa kasvumallissa mittaa kokonaistuottavuuden taso *A*. Yritän seuraavaksi perustella miksi.

1950- ja 1960-luvuilla uskottiin vahvasti, että investoiminen kiinteään pääomaan on menestymisen avain. Tämä oppi eli vahvasti kehitystaloustieteilijöiden (mm. Rostow ja Lewis) kirjoituksissa sekä poliitikkojen puheissa. Nikita Hruštševin kerrotaan sanoneen, että poliitikot rakentavat mielellään siltoja – sinnekin missä ei ole jokia. Meillä tämän pääomafundamentalismin opin lanseerasi vuonna 1952 silloinen pääministeri Urho Kekkonen kirjassaan *Onko maallamme malttia vaurastua?* Olen omassa kirjassani *Tehoton pääoma* (Pohjola 1996) näyttänyt, miten tämä malli toimi aikansa mutta johti meillä investointien tehottomuuteen, joka paljastui viimeistään 1990-luvun lamavuosien aikana.

Kuvio 9.4. Talouskasvun polarisoituminen



Kuvioissa on kuvattu vuoden 1996 dollareissa arvioitujen asukasta kohti laskettujen bruttokansantuotteiden jakaumat vuosina 1960 (ylempi kuvio) ja 1998 (alempi kuvio); havaintoina 108 kansantaloutta. Aineistolähteenä on Penn World Table 6.

On monta tapaa todistaa, miksei investoiminen kiinteään pääomaan voi selittää havaittuja kehityseroja maiden kesken. Ensimmäinen ja mielestäni nerokkain on Feyrerin (2001) havainto siitä, että pääomakannan (tai oikeastaan pääomakertoimen) jakauma maiden kesken ei ole kaksihuippuinen vaan se on yksihuippuinen. Se ei voi siten selittää elintason jakauman muotoa kuviossa 9.4. Täsmälleen sama havainto pätee myös henkiseen pääomaan eli koulutukseen. Feyrerin johtopäätöksenä onkin, että teknologia eli kokonaistuottavuuden taso on pulman ratkaisu. Se on jakautunut maiden kesken tavalla, joka selittää empiiriset havainnot.

Toinen tapa mitata investointien vaikutusta on laskea kuinka suuri ero yhtälön (2) investointiasteessa s selittää havaitut erot elintasossa y = Y/L. Tarkastellaan kahta kansantaloutta A ja B, jotka eroavat toistaan vain investointiasteiden suhteen. Yhtälöstä (2) seuraa, että

$$\frac{y_A}{y_B} = \left(\frac{s_A}{s_B}\right)^{\frac{b}{1-b}} \qquad (3)$$

Jos pääoman osuus kansantulosta b on esimerkiksi 0,3, niin investointiasteen maassa A tulee olla 1086 kertaa niin suuri kuin maassa B, jotta A:n elintaso olisi 20-kertainen maahan B verrattuna. Jos b = 0,4, niin kertoimeksi tulee 89, mutta sekin on liian suuri ollakseen realistinen. Itse asiassa monet kehitysmaat investoivat kansantulostaan suhteellisesti enemmän kuin teollisuusmaat. Suomi teki eräänlaisen investoimisen maailmanennätyksen vuosina 1950-1990, mutta vaikka elintasoeroa rikkaisiin maihin kurottiin umpeen, niin kuten kuvio 9.2 osoittaa, taso jäi kuitenkin 70 prosenttiin Yhdysvaltojen tasosta (Pohjola 1996).

Kolmas tapa perustella sitä, miksei kiinteä pääoma voi olla kehityserojen selitys, on todeta, että maailman pääomanliikkeistä 90 prosenttia tapahtuu rikkaiden maiden kesken. Jos pääoma olisi talouskasvun keskeisin selittäjä, niin investointien tuottoaste olisi korkea kehitysmaissa, ja suorat sijoitukset kohdentuisivat näihin talouksiin. Käytännössä käy kuitenkin päinvastoin: pääoma tulee pääoman luo (Easterly 2001; Easterly ja Levine 2001). Miksi? Siksi, että pääoma on komplementaarinen teknologian kanssa. Siellä missä on ideoita ja tietoa, sinne investoiminen on myös kannattavaa.

Näitä perusteluja voi käyttää myös sen todistamiseen, ettei myöskään henkinen pääoma voi olla elintasoerojen tärkein selitys. Yhtälöä (3) voidaan soveltaa, kun säästämisaste tulkitaan investoinniksi koulutukseen. Henkisen pääoman tulo-osuutena voidaan käyttää myös haarukkaa 0,3-0,4 (Mankiw,

Romer ja Weil 1992), jolloin päästään täsmälleen samoihin tuloksiin kuin edellä.

Henkiseen pääomaan pätee myös se havainto, että se keskittyy niihin maihin, joilla on sitä jo ennestään runsaasti. Miksi? Siksi, että niillä on teknologiaa – tietoa, ideoita ja osaamista – joita soveltamalla koulutettu työvoima saa investoinneilleen tuottoa. Koulutuksesta ei ole mitään hyötyä paikassa, jossa ei ole muita koulutettuja ihmisiä. Siksi koulutettu työvoima muuttaa köyhistä maista rikkaisiin, eikä päinvastoin. Rikkailla mailla on teknologia, joka köyhistä maista puuttuu. Teknologia ja henkinen pääoma ovat komplementteja – toisiaan täydentäviä.

Totuus on, että talouskasvun ja koulutusinvestointien välinen korrelaatio on negatiivinen maailmantaloudessa (Easterly 2001). Tämä pätee myös OECD-maihin 1990-luvulla. Näyttää siltä, että useat maat investoivat liikaa henkiseen pääomaan. Koulutus ei ole enää investointia vaan kulutusta. Se voi toki silti olla hyvinvoinnin keskeinen tekijä, vaikka ei nostaisikaan elintasoa.

Fyysisen ja henkisen pääoman lisäksi myös investoinnit aineettomaan pääomaan kiihdyttävät ensin talouskasvua ja näkyvät lopulta elintason nousuna. Näitä investointeja ovat tutkimus- ja kehitystoiminnan lisäksi kaikki ne yritysten aktiviteetit, jotka lisäävät nykyisen tuotannon sijasta yrityksen tuotantokapasiteettia ja siten tulevaa tuotantoa. Uusien yritysten perustaminen on tästä hyvä esimerkki. Toimintaa aloittaessaan yrittäjät investoivat yritykseensä suuria summia menetetyn palkkatulon muodossa. Yritysorganisaation kehittäminen on toinen esimerkki. Liiketoimintaprosessien kehittäminen vaatii sekä aikaa että resursseja. Sosiaalisen pääoman kehittäminen käy esimerkistä koko kansantalouden tasolla. Tämä on pääomaa, joka vähentää liiketoiminnan transaktiokustannuksia esimerkiksi siten, että ihmisten luottaessa toisiinsa sopimusten tekemiseen ja valvomiseen ei tarvitse käyttää paljon resursseja. Luottamuksen rakentaminen ja "kansallisen kulttuurin" luominen ovat kuitenkin aikaa ja resursseja vieviä prosesseja.

Aineettoman pääoman merkityksen arvioiminen on hankalaa siksi, ettei käytettävissä ole empiirisiä tietoja sen tulo-osuudesta. Kuinka suuri osa kaikesta taloudellisesta toiminnasta käytetään tulevan tuotoksen kasvattamiseen? Toisin sanoen, kuinka paljon taloudellisia uhrauksia voitaisiin vähentää nykyisen tuotantotason kärsimättä esimerkiksi karsimalla T&K-toiminnot, henkilöstöhallinto, mainoskulut, yritysten sisäinen koulutus jne? Parente ja Prescott (2000) arvelevat näiden investointien olevan suuria, mutta enintään 30-50 prosenttia Yhdysvaltojen kansantuotteesta. Käyttämällä tulo-osuutta 0,3, päästään samoihin johtopäätöksiin kuin edellä kiinteän pääoman osalta. Jos aineettoman pääoman tulo-osuus on 0,5, niin silloin 20-kertainen

ero investoinneissa aineettomaan pääomaan johtaisi pitkällä aikavälillä 20-kertaiseen eroon elintasossa. Tämä vaikuttaa uskottavalta siinä mielessä, että investointiaste aineettomaan pääomaan voi kehitysmaissa hyvinkin olla vain 1/20 Yhdysvaltojen tasosta. Silloin kuitenkin näiden investointien tuotto olisi korkea, noin 40 prosenttia, mistä pitäisi seurata se, että kehitysmaat panostavat voimakkaasti aineettomaan pääomaan. Näin ei näytä tapahtuvan.

Suomen kasvupolitiikassa tapahtui 1990-luvun laman jälkeen muutos, joka on korostanut aineettomien investointien merkitystä aineellisten ja koulutusinvestointien asemasta. Tutkimus- ja kehitysinvestointien osuus kansantuotteesta on nyt maailman kärkeä. Politiikan suunnanmuutos on monessa mielessä perusteltu, mutta edellä esitettyyn perustuen on helppo todeta, että pelkästään tällä tavoin emme pääse kolmen menestyneimmän kansakunnan joukkoon vuoteen 2010 mennessä, emmekä vielä pian sen jälkeenkään. Erot T&K-investointiasteissa ovat niin pieniä, ettei niillä ole kovin suurta vaikutusta elintasoeroihin, ellei tuotannon jousto T&K-pääoman suhteen (kerroin b yhtälössä (1)) ole huomattavasti paljon suurempi kuin pääoman muiden komponenttien suhteen.

KOKONAISTUOTTAVUUDEN MERKITYS

Tarkastellaan seuraavaksi, voisivatko erot kokonaistuottavuuden tasossa selittää havaitun suuruiset erot elintasossa. Oletetaan siksi, että maat A ja B investoivat yhtä paljon kiinteään, henkiseen ja aineettomaan pääomaan. Nyt yhtälöstä (2) seuraa, että

$$\frac{y_A}{y_B} = \left(\frac{A_A}{A_B} \right)^{\frac{1}{1-b}}. \qquad (4)$$

Useiden empiiristen kasvututkimusten mukaan (esimerkiksi Romer, Mankiw ja Weil 1992) kiinteiden ja henkisen pääoman osuus kansantulosta on yhteensä likimain 2/3. Tätä b:n arvoa käyttämällä yllä esitetystä seuraa, että esimerkiksi kolminkertaiset erot kokonaistuottavuudessa johtavat 27-kertaisiin eroihin elintasossa. Näin ollen uskottavan suuruisilla tuottavuus- tai tehokkuuseroilla voidaan selittää havaitun suuruiset erot kansantuotteessa asukasta kohti.

Tämä havainto selittää, miksi empiirisen kasvututkimuksen painopiste on viime vuosina muuttunut kansantuotteen kasvun analyysistä kokonaistuottavuuteen (total factor productivity, TFP) vaikuttavien tekijöiden tutki-

miseen. Modernin kasvututkimuksen isä Robert M. Solow totesi hiljattain seuraavasti (Solow 2001):

> "If they mean anything at all, those many right-hand-side variables in growth regressions are determinants of TFP. But then they should be selected with that function in mind, and TFP (or its growth rate) should be the left-hand-side variable."

Solow'n toteamus pätee erityisen hyvin Suomen kansantalouden kilpailukyvyn erittelyyn. Yhtälöä (4) voi nimittäin soveltaa kansantaloutemme kokonaistuottavuuden tason arvioimiseen. Kuvion 9.2 mukaan kansantuote asukasta kohti on meillä ollut noin 70 prosenttia Yhdysvaltojen tasosta jo 25 vuoden ajan. Jos nyt oletamme, että $b = 2/3$ ja että Suomi ja Yhdysvallat ovat investoineet kansantulostaan suhteellisesti yhtä paljon pääoman eri komponentteihin, niin yhtälöstä seuraa, että kokonaistuottavuuden tason pitäisi meillä olla noin 88 prosenttia amerikkalaisesta tasosta. Kansainvälisissä vertailuissa se on kuitenkin mitattu selvästi alhaisemmaksi.

Islam (2000) arvioi kokonaistuottavuuden olleen Suomessa noin 60 prosenttia Yhdysvaltojen tasosta 1970-luvulla ja vain liki 50 prosenttia 1980-luvulla. Sijoituksemme lähes 100 maan vertailussa olivat vastaavasti 20:s ja 23:s. Hall ja Jones (1999) päätyvät paljon myönteisempää kuvaan: 80 prosenttia vuonna 1988. Itse olen laskenut sen olleen 74 prosenttia vuonna 1990 (Pohjola 1996).

Kokonaistuottavuus voi meillä olla alempi kuin suhteellinen tulotaso, jos olemme investoineet kansantuotteestamme vertailumaata enemmän. Näinhän meillä tehtiin kiinteän pääoman osalta 1950-luvulta aina 1980-luvun lopulle saakka. Investointimme koulutukseen ja tutkimukseen ovat myös kansainvälisesti huipputasoa. Edellä esitetyn perusteella näiden vaikutus elintasoon ei kuitenkaan voi olla niin suuri kuin poliittiset päättäjät uskovat. Kokonaistuottavuuteen vaikuttaviin tekijöihin olisi siksi kiinnitettävä nykyistä enemmän huomiota.

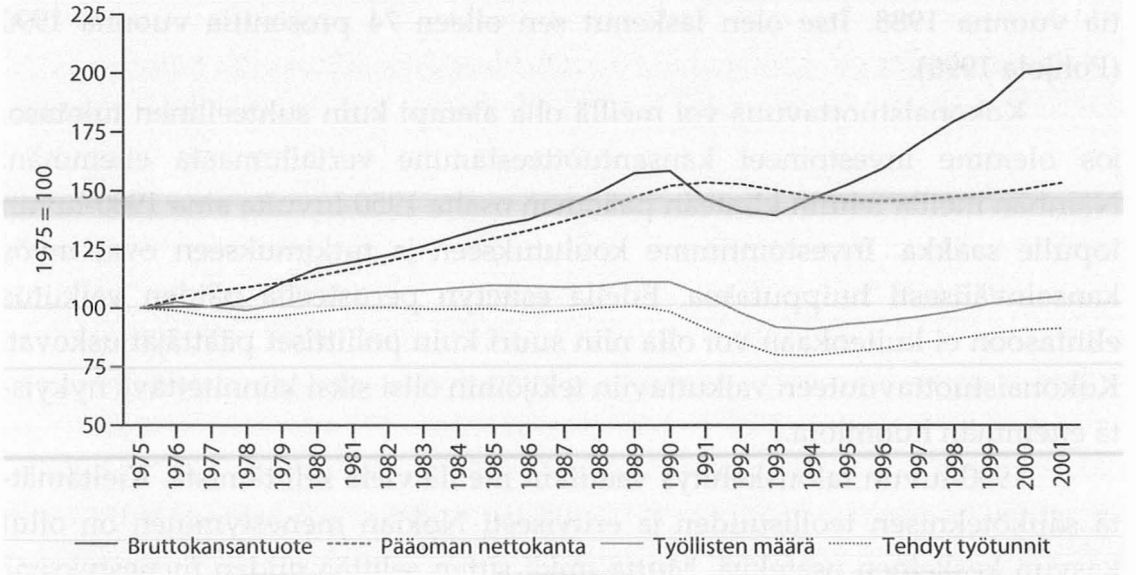
1990-luvun talouskehitys vaatiikin meillä vielä selittämistä. Kieltämättä sähköteknisen teollisuuden ja erityisesti Nokian menestyminen on ollut kasvun keskeinen osatekijä. Mutta mikä sitten selittää niiden menestyksen? Nokiahan perustettiin jo vuonna 1865. Nopeaa toipumista lamasta ei myöskään voi panna yksin tutkimus- ja kehitystoiminnan tiliin, koska sen kansantuoteosuus on kasvanut maailman kärkiluokkaan vasta 1990-luvun jälkipuoliskolla.

Kuvio 9.5 näyttää, miten kansantaloutemme toiminta tehostui viime vuosikymmenellä. Tuotannon määrä kasvoi 29 prosenttia vuodesta 1990 vuo-

teen 2001, mutta työpanoksen määrä ei ole kasvanut työtunneilla eikä työntekijöiden määrällä arvioituna. Kaikkein suurin muutos on kuitenkin tapahtunut pääoman käytössä. Reaalinen nettopääomakanta kasvoi lähes samaa tahtia tuotannon kanssa aina lamakauteen asti, mutta sen jälkeen kasvua ei ole tapahtunut. Tehoton pääoma on muuttunut tehokkaaksi. Toteutunut talouskehitys todistaakin mielestäni oikeaksi *Tehoton pääoma* -kirjani perusteesin. 1980-luvulla tapahtunut pääomanliikkeiden ja rahamarkkinoiden säänte-lyn purkaminen sekä talouden avautuminen kansainväliselle kilpailulle toivat elinkeinoelämässämme piilleet tehottomuudet esiin. Rakenteellisten ongelmien korjaamiseen ryhdyttiin kuitenkin vasta taloutemme kriisin pakottamana, eikä tämä prosessi ole vieläkään päättynyt.

Tämä analyysi sai aikanaan voimakasta kritiikkiä – sekä asiallista että asiatonta. Jälkimmäiseen luokkaan kuuluvat Matti Virénin (1996) kirja-arvostelu sekä Heikki Urmaksen (1998) mielipiteet kirjassaan *Uhrina työtön*, jossa hän julkaisi Helsingin kauppakorkeakoulussa 1990-luvulla vaikuttaneen ns. kolmen tohtorin (Veikko Jääskeläinen, Heikki Urmas ja Fedi Vaivio) talouspoliittisen ohjelman. Toteutuneen talouskehityksen valossa olisikin näi-

Kuvio 9.5. Kansantuote, pääomakanta ja työpanos (1975-2001)



Kuvion käyrät kuvaavat indeksisarjoja, 1975 = 100. Asuinrakennukset eivät ole mukana pääomakannassa eivätkä asuntojen omistus ja vuokraus bruttokansantuotteessa. Lähteinä ovat pääomakannan osalta Aulin-Ahmavaara ja Jalava (2002) ja muiden sarjojen osalta kansantalouden tilinpito- ja työllisyystilastot.

den arvostelijoiden nyt syytä vastata kysymykseen, mikä muu tekijä kuin menneiden investointien tehottomuus voi selittää kansantaloutemme pääomakannan kehityksen 1990-luvulla.

Kuvion 9.5 perusteella voi päätellä kansantaloutemme toiminnan tehostuneen ja kokonaistuottavuuden nousseen aivan uudelle tasolle. Tämä merkitsee sitä, että yhtälöissä (1) ja (2) esitetyn kasvuteorian tehokkuusparametrin A arvo on kasvanut. Edellä esitetyn mukaan tämä parametri on tärkein maiden välisiä elintasoeroja selittävä tekijä. Teorian perusteella voi siis päätellä Suomella olevan edessään parempi tulevaisuus. Mutta mitkä tekijät ovat muutoksen saaneet aikaan, ja miten näitä tekijöitä voisi talouspolitiikalla entisestään edistää?

KANSALLISEN KILPAILUKYVYN KANNALTA KESKEISIÄ TEKIJÖITÄ

Seuraavat tekijät ovat mielestäni keskeisiä mietittäessä kansantaloutemme menestymiseen vaikuttavia tekijöitä:

Tuottavuuden ja tehokkuuden esteet muilla kuin teollisilla toimialoilla

Työn tuottavuus (bruttokansantuote/työtunnit) on Suomen kansantaloudessa OECD-maiden keskimääräistä tasoa (Pohjola 2003). Tiedämme kuitenkin, että teollisuudessa työn tuottavuus on meillä huippuluokkaa. Olemme saavuttaneet – ja jopa ohittaneet – sekä monet vanhat teollisuusmaat että Yhdysvallat. Tuottavuus nousi rajusti laman aikana tunnetusta syystä: heikoimmat yritykset poistuivat markkinoilta ja jäljelle jäivät vain vahvimmat. Työn määrä väheni voimakkaasti. Laman jälkeen tuottavuuden kasvuvauhti on säilynyt nopeana lähinnä elektroniikkateollisuuden menestymisen myötä, mikä on ylläpitänyt korkeaa tasoa.

Kaikesta tästä seuraa kuitenkin, että työn tuottavuuden on oltava selvästi keskimääräistä alempi muilla kuin teollisilla toimialoilla. Mutta kuinka paljon se sitä on ja mistä syystä? Ensimmäisiä mittauksia on jo tehty (Mankinen, Rouvinen ja Ylä-Anttila 2002), mutta syitä ei ole juurikaan selvitetty.

Syyt eivät voi liittyä koulutukseen eivätkä investointeihin, sillä nämä samat tekijät vaikuttavat myös teollisilla aloilla. Syyn on siksi löydyttävä tehokkuuden ja kokonaistuottavuuden tasosta. Näyttää siltä, että syystä tai toisesta suomalaiset palvelualan yritykset eivät käytä parasta mahdollista teknologiaa. Koska tällainen teknologia on – ainakin periaatteessa – vapaasti saatavilla, niin sen omaksumiselle on oltava käytännön esteitä. Asia voi olla

yksinkertaisesti niin, että yrityksillä ei ole kannusteita tai pakkoa sen omaksumiseen. Kilpailun vähyys voi olla mahdollinen selitys.

Taloudellisen toiminnan alueellinen keskittyminen

Kuviossa 9.4 näytettiin miten maiden väliset elintasoerot ovat kasvussa. Aivan sama ilmiö tapahtuu maan sisällä eri alueiden kesken ja täsmälleen samasta syystä: teknologia, aineellinen ja henkinen pääoman keskittyvät alueellisesti. Talouspolitiikan olisi siksi oltava näiden asioiden osalta keskenään ristiriidatonta. Ei voi olla niin, että yhtäältä pyrimme nostamaan yhteiskuntamme elintasoa ja samalla estämään alueiden kesken tapahtuvaa rakennemuutosta. Jos näin tehdään, niin aluepolitiikkaa on harjoitettava muihin tavoitteisiin kuin elintason nostamiseen pyrkien. Tätä keskustelua ei Suomessa käydä eikä vaihtoehtoja pohdita.

Tieto- ja viestintätekniikan käytön vaikutukset

Suomi on yksi johtavista tietoyhteiskunnista, mutta olemme toistaiseksi hyötyneet tieto- ja viestintekniikan vallankumouksesta pääosin tavaroiden ja palvelujen tuottajana. Uuden teknologian vaikutus tuottavuuteen ei ole talouden muilla sektoreilla ollut niin suuri kuin se voisi olla, kun otamme huomioon hyvän infrastruktuurimme (Koski, Rouvinen ja Ylä-Anttila 2002, Pohjola 2003).

Palvelualojen tuottavuuden parantaminen on avainasemassa uutta teknologiaa hyödyntävässä taloudessa. Teknologia ei kuitenkaan yksin ratkaise mitään ongelmaa. Tietokoneiden kapasiteetin kasvu ja hintojen lasku ovat vain voimia, joka tekevät toimintatapojen muutokset mahdollisiksi. Liiketoiminnassa tämä tarkoittaa uusien tuotteiden ja palvelujen kehittämistä sekä uusien organisaatiomuotojen (johtamisjärjestelmien, kannusteiden ym.) omaksumista – siis informaation tehokasta hyödyntämistä. Näiden tekijöiden sovittaminen yhteen ei ole helppo ongelma. Liiketoimintaosaaminen on avainasemassa. Se on meillä kuitenkin jäänyt teknologisen osaamisen varjoon sekä tutkimuksessa että talouspoliittisessa päätöksenteossa.

T&K-investointien vaikuttavuus

Suomi on tekemässä jälleen investoimisen maailmanennätystä – tällä kertaa tutkimus- ja kehitysinvestoinnissa. Viime vuosikymmenten kokemuksiimme vedoten on helppo jo nyt ennustaa, etteivät tulokset tule olemaan ennätys-

luokkaa. Tehottomuuden estämiseksi olisi kiireesti ryhdyttävä mittaamaan näiden investointien vaikuttavuutta eli kontribuutiota kansantuotteen kasvuun. Eräät maailman johtavat tuottavuustutkijat tekevät tällaista tutkimusta seitsemän OECD-maan osalta, mutta Suomi ei valitettavasti ole mukana (Mohnen, Mairesse ja Dagenais 2001). Ideana on soveltaa T&K-investointeihin samantapaista kokonaistuottavuuden laskentakehikkoa kuin talouskasvun analyysiin.

Verotus ja työpanos

Kansainvälisistä vertailuista tiedämme, että työtuntien määrällä asukasta kohti mitattu työpanos on meillä myös OECD-maiden keskimääräistä tasoa ja että se on voimakkaasti vähentynyt 1990-luvulla (Pohjola 2003). Julkisessa keskustelussa verotuksen korkeaa tasoa on pidetty sekä tämän kehityksen mahdollisena selityksenä, että talouden tulevan kasvun esteenä. Yllä esitettyä kasvuteorian kehikkoa voi käyttää myös verotuksen aiheuttaman kannustinvaikutuksen arvioimiseen. Prescott (2002) on arvioinut työn verotuksen olevat keskeisin Euroopan ja Yhdysvaltojen elintasoeroa selittävä tekijä. Suomi tosin poikkeaa monista muista Euroopan maista korkean työttömyyden vuoksi, jolloin työpanoksen vähäisyys Yhdysvaltoihin verrattuna johtuu ehkä muista syistä kuin työn tarjonnan vapaaehtoisesta vähentämisestä raskaan verotuksen takia. Verotuksen pitkän ajan kannustinvaikutukset kaipaavat nykyistä perusteellisempaa pohdintaa.

KIRJALLISUUTTA

Aulin-Ahmavaara, P. ja Jalava, J. (2002): "Pääomapanos ja sen tuottavuus Suomessa vuosina 1975-2001", VATT keskustelualoitteita no. 294.

Easterly, W. (2001): *The Elusive Quest for Growth*, MIT Press, Cambridge, Massachusetts.

Easterly, W. ja Levine, R. (2001): "It's not factor accumulation: Stylized facts and growth models", *The World Bank Economic Review* 15: 177-219.

Feyrer, J. (2001): "Convergence by parts", julkaisematon, Dartmouth College.

Hall, R. ja Jones , C. (1999): "Why do some countries produce so much more output per capita than others?", *Quarterly Journal of Economics* 114: 83-116.

Islam, N. (2000): "Productivity dynamics in a large sample of countries", julkaisematon, Emory University.

Kekkonen, U. (1952): *Onko maallamme malttia vaurastua?*, Otava, Helsinki.

Koski, H., Rouvinen, P. ja Ylä-Anttila, P. (2002): *Tieto & talous. Mitä uudesta taloudesta jäi*, Edita Prima Oy, Helsinki.

Krugman, P. (1994): *The Age of Diminished Expectations*, MIT Press.

Maddison, A. (1995): *Monitoring the World Economy*, OECD, Paris.

Mankinen, R., Rouvinen, P. ja Ylä-Anttila, P. (2002): "Palveluiden tuottavuus—kilpailu ja teknologia muuttavat rakenteita", ETLAn keskusteluaiheita no. 829.

Mankiw, N.G., Romer, D. ja Weil, D.N. (1992): "A contribution to the empirics of growth", *Quarterly Journal of Economics* 109: 407 - 437.

Mohnen, P., Mairesse, J. ja Dagenais, M. (2001): "The importance of being innovative: A comparison of seven OECD countries", julkaisematon käsikirjoitus.

Parente, S.L. ja Prescott, E.C. (2000): *Barriers to Riches*, MIT Press, Cambridge, Massachusetts.

Pohjola, M. (1996): *Tehoton pääoma*, WSOY, Porvoo.

Pohjola, M. (2003): "Elintaso, tuottavuus ja kilpailukyky uudessa taloudessa", teoksessa Pekkarinen, J. ja Kavonius, I.K. (toim): *Suomi, EU ja maailma*, Työväen sivistysliitto, Helsinki, 64-74.

Porter, M.E. (1990): "The competitive advantage of nations", *Harvard Business Review* 1990: 73 - 93.

Prescott, E.C. (2002): "Prosperity and depression", *American Economic Review, Papers and Proceedings* 92, 1-15.

Rouvinen, P. (2001): "Finland on top of the competitiveness game?", *The Finnish Economy and Society* 4/2001, ETLA.

Rouvinen, P. ja Vartia, P. (2002): "Suomi—maailman kilpailukykyisin?", *Suhdanne* 1/2002, 93-100.

Solow, R.M. (2001): "Applying growth theory across countries", *The World Bank Economic Review* 15: 283-288.

UNDP (2001): *Human Development Report 2001. Making New Technologies work for Human Development*, United Nations Development Programme, Oxford University Press.

Urmas, H. (1998): *Uhrina työtön*, Gummerus Kirjapaino Oy, Saarijärvi.

Virén, M. (1996): "Suomen talouden ongelmat—uutta näkökulmaa etsimässä", *Kansantaloudellinen aikakauskirja* 92, 483-489.

10. Kasvun selitykset ja Suomen talouskasvu

Paavo Okko*

Professori Pentti Vartia on ollut aina kiinnostunut suurista taloudellisista kysymyksistä. Yksi niistä on taloudellinen kasvu, joka ratkaisee myös suomalaisen yhteiskunnan menestyksen. Taloudellisen kasvun ikuista kysymystä on viime aikoina tarkasteltu uudella tavalla. Kansantaloustieteellisen kasvuteorian uusi aalto alkoi puhua 1980-luvulla kasvun endogeenisuudesta. Kasvuvauhti on tämän ajattelun mukaan ainakin jossakin määrin omissa käsissämme. Suomalaiset ovat tietysti uskoneet tähän jo pitkään. Kiinnostava yhteensattuma onkin, että samoihin aikoihin kun kasvuteoria kehittyi tähän suuntaan, Suomessa tapahtui sellaista, joka vahvistaa uskoa teorian toimivuuteen. Tältä kannalta tärkeä ajankohtainen kysymys on, mitä Suomen tulevasta kasvusta voisi päätellä uuden kasvuteorian valossa.

Tarkastelen tässä kirjoituksessa kasvun selityksiä Suomen esimerkin valossa. Haluan korostaa institutionaalisten ratkaisujen merkitystä. Suomen kyky kuroa umpeen korkeammalla tulotasolla olevien maiden etumatkaa on ollut vahva. Laman jälkeen tapahtunut nousu on perustunut talouden innovaatiotuotantosektorin nopeaan kasvuun. Kasvun kiihtyminen tämän prosessin aikana on teorian mukainen tulos, joka ei kuitenkaan merkitse pysyvää muutosta kasvuvauhdissa, vaikka panostukset kestäisivätkin korkeammalla tasolla. Suomella on nyt uudenlainen kasvuhaaste edessään.

KASVUN YLEISET SÄÄNNÖNMUKAISUUDET

Kansantulon tasoa ja sen kasvua koskevina yleisinä empiirisinä havaintoina voidaan esittää seuraavat "stilisoidut faktat" (esim. Jones 1998, 5-16):

1. Henkeä kohti lasketun tulon taso vaihtelee maitten välillä suuresti.
2. Eri maiden kasvuvauhti vaihtelee olennaisesti.
3. Talouksien kasvuvauhdit eivät pysy ajan myötä vakioina.
4. Maan asema tulotasoasteikossa voi liikkua sekä ylös- että alaspäin.

* Paavo Okko on professori Turun kauppakorkeakoulussa.

5. Erityisesti Yhdysvaltoja koskevana tosiasiana on esitetty: pääoman reaalituoton taso on pitkällä tähtäyksellä vakio; funktionaalinen tulonjako on samoin vakaa sekä henkeä kohti lasketun tulon kasvuvauhti on ollut pitkän päälle melko vakaa.
6. Tuotannon kasvun ja kansainvälisen kaupan kasvu liittyvät läheisesti toisiinsa.
7. Sekä paljon että vähän koulutettu työvoima liikkuu kohti korkean tulotason alueita ja maita.

Maailman maiden tulotasoerojen tarkastelu osoittaa todellakin valtavia eroja. Esimerkiksi Kanadan kansantulo on 107 kertaa korkeampi kuin Burundin (Easterley – Levine 2003, 4). Kun talousteoria pyrkii selittämään tällaisia eroja, se joutuu ottamaan lähtökohdaksi paljon laveamman näkökulman kuin mitä tuotantofunktion osoittama panoksien ja tuotoksen välinen yhteys tarjoaa. Silloin on hyvä kysyä, mikä johtuu maantieteellisestä sijainnista ja luonnon varoista, mikä instituutioista ja mikä politiikan eroista. Näitä asioita on tarpeen miettiä, kun yritetään ymmärtää kasvuvauhdin vaihtelun perimmäisiä syitä.

Suomen tulotaso oli itsenäistymisen aikoihin noin puolet Iso-Britannian tasosta (Vartia – Ylä-Anttila 1996, 63). Nyt olemme EU:n ja Iso-Britannian tason yläpuolella. Suomi on pystynyt suhteellisen nopeaan kasvuun ja ollut maa, joka on pystynyt kuromaan umpeen rikkaampien maiden etumatkaa. Ns. konvergenssitutkimus osoittaa, että pääasiallinen havainto on Euroopassa yleensäkin ollut alueellinen tulotasojen lähentyminen. Sen vauhti on kuitenkin vaihdellut ja se on mieluummin hidastunut viime aikoina (ks. esim. Sala-i-Martin 1994 ja Puga 2001). Melko samoin voidaan sanoa Suomen sisäisistä alueiden välisistä keskimääräisistä tulotasoista (ks. esim. Kangasharju 1998). Tuotannon uudelleensijoittumisen ja väestön muuton seurauksena eri alueiden henkeä kohti lasketut tulot lähestyvät toisiaan, vaikka tuotantorakenteet erilaistuvat ja väestö keskittyy. Mielenkiintoinen havainto on myös se, että EU-maiden kesken erot ovat supistuneet, mutta maiden sisällä tulotasoerot ovat 1990-luvulla mieluummin laajentuneet (Puga 2001, 4 ja Giannetti 2002). Integraation ja kasvun myötä tapahtuu siis sekä konvergoitumista että divergoitumista. Tulotasojen lähentyminen vaatii sitä, että alueet ja maat pääsevät hyödyntämään samaa teknologiaa. Siksi tulotasojen lähestyminen ei ole universaali ilmiö teknologisen kilpailun maailmassa. Selitystä jää kaipaamaan myös se, miksi maat eivät ole päässeet käyttämään samaa teknologiaa.

Suomessa pääoman tuottoaste sekä tuottavuus olivat viimeaikaista tasoa alempana ennen lamaa vallinneen korkean investointiasteen aikana (Aulin-Ahmavaara – Jalava 2003). Korkosäännöstelyn aikana verojen jälkeinen investointien rahoituskustannus oli aika ajoin jopa negatiivinen. Investointien tehokkuus oli heikko kasvuvauhdin näkökulmasta (Pohjola 1996). Laman jälkeen kiinteiden investointien taso on ollut olennaisesti aikaisempaa alempi. Kiinteän pääoman muodostuksen tahti on ollut nyt jatkuvasti vajaa 20% BKT:sta, kun se aikaisemmin oli miltei aina tuon tason yläpuolella.

Samanaikaisesti on käynyt niin, että funktionaalinen tulonjako on poikennut aikaisemmasta käyttäytymisestään. Palkkasuhde nousi lamassa tyypilliseen tapaan, mutta laman jälkeisen nousun aikana se laski aikaisempaa trenditasoaan alemmaksi. Samalla pääoman tuottoaste on ollut aikaisempaa korkeammalla. Erityisen vahva muutos on tapahtunut elektroniikkateollisuudessa. Kasvuempirian kannalta voi arvella, että nykyinen taso ei ole uusi pitkän aikavälin tasapainotaso vaan poikkeama, joka vähitellen väistyy. Työllisyysasteen nykyinen alhaisuus saman ilmiö yksi puoli. Sen toivottu korjaantuminen tulee olemaan osa tätä tasapainottumista.

Kasvun tosiasioista varmaan parhaiten Suomessa sisäistetty osa on viennin merkitys. Sekä suhdanne- että kasvutekijänä vienti on ollut keskeisessä asemassa. Viennillä on ollut aikaisemmin hyvin kiinteä yhteys maan luonnonvaraperustaan. Nyt on korostunut osaamispanoksen merkitys maan vientipohjan vahvistajana.

Suomesta on muutettu korkeamman tulotason Ruotsiin, ja maan sisällä on seurattu samaa periaatetta. Työvoima on liikkunut kasvun myötä keskusseuduille, joissa työttömyys on alempi ja tulotaso korkeampi kuin lähtöseuduilla. Samalla on käynyt niin, että inhimillinen pääoma on liikkunut sinne, missä sitä entuudestaan on jo paljon (Ritsilä 2001). Vaikka muuttoliike eriyttää alueiden tuotantokykyä, se ei kuitenkaan lisää alueiden henkeä kohti laskettujen keskimääräisten tulojen eroja, vaan tulotasojen konvergenssi on ollut pääsääntö Suomessakin.

Suomi on ollut yksi ns. kiinnikuromishypoteesin mukaisesti edennyt maa (ks. Vartia – Ylä-Anttila 1996, 63). Maa on noussut Euroopan periferiasta Euroopan rikkaiden maiden keskiarvon yläpuolelle. Tämä merkitsee sitä, että kehitys on selitettävissä kasvuteorian säännönmukaisuuksilla, vaikka tapauksemme kuuluukin havaintojen yhteen erityisryhmään. Seuraavaksi tarkastellaan kasvua pohtimalla, mikä merkitys asiassa on resurssiperustalla, instituutioilla ja politiikalla. Näiden tekijöiden avulla on kasvututkimuksessa haettu eräänlaisia perimmäisiä selityksiä. Tämän jälkeen tarkastellaan vielä

Suomen nykyistä kasvuhaastetta rajatummin osaamispanostuksen näkökulmasta.

RESURSSIPERUSTA, INSTITUUTIOT JA POLITIIKKA

Pohdittaessa sitä, miksi eri yhteiskunnissa taloudellisen vaurauden taso vaihtelee niin suuresti, joudutaan miettimään, miksi tuotannontekijät ovat erilaiset eri maissa. Tuotantofunktiolla kuvatut taloudellisen kasvun välittömät selittäjät kaipaavat selitystä. Tällaista näkökulman avartamista voidaan tehdä monella tavalla. Eräs jäsentely perustuu maantieteellisten resurssitekijöiden, instituutioiden ja harjoitetun politiikan käyttämiseen selittäjinä (Easterley – Levine 2003). Usein myös selitetään talouden kasvua ihmisten kulttuurisiin ja henkisiin ominaisuuksiin viittaamalla, jolloin esimerkiksi uskonto näyttäytyy kasvun selittäjänä (ks. esim. Landes 1998). Ihmisten ominaispiirteisiin perustuvia selityksiä on käytetty ahkerasti historiassa. Nykyisin niistä moni näyttää melko rasistiselta.

Maantieteellisestä sijainnista ja luonnonvaroista lähtevä resurssiperustaselitys katsoo, että ympäristö vaikuttaa suoraan luonnonvarojen sekä työvoiman ja teknisen tehokkuuden vaihteluun. Sitä siis voidaan kuvatakin perinteisen tuotantofunktion avulla. Sijainti lauhkealla vyöhykkeellä tai tropiikissa on hyvä esimerkki. Eron merkitys voi kuitenkin muuttua historian kulussa. Kanadassa per capita -tulo on nyt siis 107 kertaa korkeampi kuin Burundissa. Siirtomaatalouden kehittymisen aikaan Burundi näytti kuitenkin paljon houkuttelevammalta kuin Kanada. Britannian ja Ranskan välisissä rauhanneuvotteluissa väitetään olleen brittien harkinnassa vaihtaa Kanada Guodeloupen saareen (Easterley – Levine 2003). Olihan Kanada filosofi Voltairen mielestäkin oikeastaan vain "muutama hehtaari lunta". Sijainnilliset vakiot voivatkin siis muuttua ajan myötä.

Instituutioiden merkitys kasvun selittäjä on korostunut endogeenisen kasvun teorian myötä. Niissäkin voivat olla taustalla ympäristötekijät, mutta koska instituutiot ovat hyvin pysyviä, niistä muodostuu oma tuotannontekijän kaltainen voimavaransa. Kehitysmaiden historiassa siirtomaatalous on ollut tekijä, joka on vaikuttanut myös instituutioiden kehitykseen. Näin luonnonvarat ja sijainti ovat voineet vaikuttaa myös instituutioihin. Kaikkiaan siis talouksien käyttämä teknologia on vahvasti endogeeninen ilmiö, johon ovat vaikuttaneet yhteiskuntien soveltamat instituutiot.

Instituutioihin liittyvänä erityisnäkökulmana viime aikoina on alettu korostaa sosiaalisen pääoman merkitystä myös taloudellisen kasvun tekijänä (ks. esim. Ruuskanen 2001). Yksilötasoisen inhimillisen pääoman tavoin sosi-

aalinen pääoma on tekijä, jota suorastaan tarvitaan ilmiöiden selittämiseen mutta jota on vaikea mitata. Sosiaalinen pääoma voidaan nähdä yhteistyötä ja koordinaatiota edistäväksi voimaksi keskinäiseen kilpailuun perustuvassa taloudessa. Sen tuotantoa edistävä vaikutus on helpompi mieltää kuvittelemalla, mitä tapahtuisi, jos se puuttuisi. Sosiaalinen pääoma edistää kasvua, mutta kasvu myös rapauttaa sosiaalista pääoma, koska se aiheuttaa rakennemuutosta ja rikkoo entisiä suhteita (ks. Routledge – von Amsberg 2003). Joskus on todettu, että sosiaalinen pääoma on kuin liima, joka pitää rakenteita koossa mutta voi olla joskus myös muutoksen este.

Harjoitettu talous- ja yhteiskuntapolitiikka voi myös vaikuttaa taloudelliseen kasvuun. Tuskin kuitenkaan on helppo esimerkiksi osoittaa, mikä on onnistuneen suhdannepolitiikan osuus keskimääräisestä kasvusta. Erilaiset linjaratkaisut synnyttävät ilmeisemmin eroja ja niitä voidaan myös mallintaa ekonometrisesti. Avotaloudellinen suuntautuminen ja osallistuminen taloudelliseen integraatioon sekä rahoitusmarkkinoita koskevat politiikkavalinnat edustavat näitä kasvun selityksiä. Myös omaksuttu valuuttakurssiregiimi ja siitä seuraava kilpailukykyvaikutus voi olla kasvun yksi selittäjä.

Tehdyssä tutkimuksessa, joka koski 72 entistä siirtomaata, tuli tulokseksi instituutioiden merkitystä korostavia johtopäätöksiä (Easterley – Levine 2003, 35-37). Maantieteellinen ympäristö näyttää vaikuttaneen instituutioiden muodostumiseen ja sitä kautta tulotasoon, mutta se ei selitä tulotason vaihtelua yli sen, mitä instituutioiden avulla voidaan selittää. Niinpä Burundin tulotaso olisi mallin mukaan 38 kertaa korkeampi, jos sillä olisi ollut Kanadan maantieteellinen ympäristö. Sen seurauksena sille olisi nimittäin muodostunut tehokkaampi institutionaalinen rakenne ja korkeampi tulotaso henkeä kohti. Talouspolitiikkakaan ei näyttänyt selittävän instituutioiden vakioinnin jälkeen tulotasoissa havaittuja eroja. On kuitenkin olemassa paljon tutkimustuloksia, joiden mukaan talouspolitiikka selittää kasvuvauhtia. Tutkijoiden mukaan heidän tulotasoja koskeneet selityksensä ovat sopusoinnussa kasvuvauhtiselitysten kanssa tulkitsemalla, että politiikan vaihtelusta seuraavalla kasvuvauhdin vaihtelulla ei tarvitse olla merkittäviä keskimääräistä tuloa muuttavia vaikutuksia.

Suomi ei siis ollut havaintoaineistossa mukana. Tuloksien heijastaminen meihin on kuitenkin kiinnostavaa vaikkakin ehkä epävarmalla pohjalla. Suomella on luonnonmaantieteellinen asema (mm. kaukana päiväntasaajasta, merkittäviä luonnonvaroja, rannikkosijainti), joka on tukenut toimivien instituutioiden syntyä (mm. äänioikeus ja toimiva demokratia, poliittinen vakaus, tehokas julkinen hallinto, vähäinen julkinen sääntely, vapaus korruptiosta ja

lahjonnasta). Tältä pohjalta voi siis odottaakin korkeata tulotasoa. Instituutioiden kehitykseen on vaikuttanut kuitenkin moni muukin seikka kuin maantieteellinen sijainti. Viron ja Karjalan maantieteellinen sijainti on tässä mielessä sama kuin Suomen. Muista kuin maantieteellisistä syistä johtuneet erilaiset institutionaaliset historiat näyttävät hyvin luontevilta selityksiltä nykyisille tuloeroille, mikä on sinänsä sopusoinnussa mainitun tutkimuksen tuloksien kanssa: instituutiot ovat avainasemassa.

Politiikkavalinnoilla ei tutkimuksen mukaan ollut kuitenkaan olennaista merkitystä yli institutionaalisen rakenteen. Tässä kohdin näyttää luontevalta väittää, että Suomen tulotason selittämiseen edellä esitetty ei vielä riitä. Erityisesti on syytä korostaa, että kansainväliseen työnjakoon osallistumisen suunnassa tehdyt ratkaisut ovat merkinneet paljon. Euroopan integraatioon osallistuneiden maiden on arvioitu saaneen integraatiosta vuotuiseen kasvuvauhtiinsa 0,6-0,8 prosenttiyksikön lisäyksen (Henrekson – Torstensson – Torstensson 1997). Tämä näkökulma täytyy olla merkittävä myös Suomen tapauksessa.

TIETOVARANTO, INNOVAATIOKYKY JA KASVU

Suomen viimeaikaista talouskasvua näyttää olevan hyvin luontevaa selittää endogeenisen kasvun malleilla. Edellä tarkastellun laajemman näkökulman valossa pitäisi tietysti kysyä, miksi juuri Suomi onnistui nostamaan innovaatiokapasiteettiaan ja käynnistämään laman jälkeen osaamisperusteisen nousun. Siihenkin on institutionaalinen ympäristömme vaikuttanut myönteisellä tavalla. Tarkastellaan seuraavaksi kuitenkin lähemmin viime aikojen kasvua endogeenisen kasvun idean avulla (kirjoittaja on myös aikaisemmin tarkastelut näitä kysymyksiä: Okko 2002).

Tyypillisissä endogeenisen kasvun malleissa on selitys talouden innovaatiokyvylle sekä tuotokselle. Niissä on kaksi sektoria: hyödykkeitä ja palveluksia tuottava sektori sekä inhimillistä pääomaa tuottava "yliopistosektori". Malleja on muodostettu usealla tavalla, mutta kriittinen tekijä, joka synnyttää sisäsyntyisen kasvun on tiedon tuotannon käsitteleminen omana tuotantoprosessinaan, josta seuraa vaikutuksia tuotantoon. Talouden innovaatiokyky on siis tuotannontekijä, mikä on sinänsä hyvin luonteva näkemys.

Talouden rakenne voidaan tällaisessa tapauksessa kuvata esimerkiksi näin (Romer 1990 ja Jones 1998, 152-154, ks. myös Okko 2000):

Tuotantofunktio	$Y = K^{\alpha}(AL_Y)^{1-\alpha}$
Pääoman kasvu	$\dot{K} = sY - dK$

Tiedon kasvu $\dot{A} = \delta L_A^{\lambda} A^{\phi}$

Työvoima on kahdessa käytössä $L = L_A + L_Y$

Pääomakannan kasvu ($\dot{K}$) määräytyy perinteiseen tapaan investointien (sY) ja pääomakannan taloudellisen poistuman (dK) erotuksena. Jalostusarvon (Y) tuottamisessa vallitsevat vakioiset mittakaavatuotot tuotannollisen työvoiman (L_Y) ja pääoman (K) suhteen, mutta siinä on mittakaavaetuja, jos tiedon taso nousee. Samalla kasvuvauhti ja sen endogeenisuus riippuvat tiedon tason noususta.

Innovaatioiden tuottamisen vauhti ($\dot{A}$) riippuu puolestaan siitä, kuinka suuri osa työvoimasta toimii tiedon tuotannossa (L_A) ja kuinka paljon heillä on käytettävissään kumuloitunutta tietovarantoa (A). Viimeksi mainitun tekijän osalta vaikutus uusien ideoiden syntyyn voi olla kahtalainen. Laaja ideavaranto voi edistää uusien ideoiden syntymistä ($\phi > 0$), eli keksijät pääsevät katsomaan eteenpäin edellisten keksijäsukupolvien hartioilta, kuten Isaac Newton asian ilmaisi. Myös toisinpäin voi argumentoida: helpommin keksittävät ideat on jo keksitty ja uudet ovat entistä vaikeampia ($\phi < 0$). Toinen tekijä on tutkimustyön tuottavuus (λ) ideoiden tuottamisessa. Tutkimustyöpanoksen lisääntyminen ei ehkä nosta samassa suhteessa innovaatioiden tuotosta ($\lambda < 1$). Tutkijoiden määrän lisääntyessä osa tuloksista ei ole enää uusia koko taloudelle, vaan tapahtuu päällekkäistä tutkimusta. Tungos lisääntyy, ja kaikki joutuvat nousemaan varpailleen, eikä kukaan näe entistä paremmin.

Investoinnit ja inhimillinen pääoma kytkeytyvät uuden näkemyksen mukaan toisiinsa niin, että investointien nousu voi nostaakin pysyvästi kasvuvauhtia. Laajenevassa pääomakannassa voidaan soveltaa kasvavaa tietovarantoa. Teorian mukaan myös kansainvälisessä integraatiossa hyötyjä syntyy juuri siitä, että entistä laajemmat idea- ja hyödykesektorit voivat olla vuorovaikutuksessa keskenään. Nykyaikaisen tietoteollisen tuotantomuodon peruskysymys on siinä, kuinka ideat materialisoidaan tuotteisiin.

Parametrien arvoista ei ole empiiristä yksimielisyyttä. On kuitenkin luontevaa ajatella, että aikaisempi tietovaranto lisää innovaatiopotentiaalia ja tutkimustyövoiman määrän lisääminen nostaa ainakin johonkin määrään saakka innovaatiotuotosta. Tällöin panoksien ja tuotannon suhteen vallitsee mittakaavaetuja, ja innovaatiokyvyn lisääminen on mahdollista omilla panostuksilla. Parametrien arvoista riippuen malli tuottaa tasapainoiseksi kasvuvauhdiksi erilaisia tuloksia. Yksinkertaisessa tapauksessa ($\lambda = 1, \phi = 0$) tietovarannon sekä kansantulon tasapainokasvu on tietosektorilla työskentelevien määrän kasvun suuruinen. Tulos muistuttaa uusklassisen teorian tasapaino-

kasvua, joka on työvoiman kasvun suuruinen, jos teknistä kehitystä ei oteta huomioon.

Suomessa tehtiin tutkimus- ja kehitystoiminnassa tutkimustyövuosia runsas 53 000 vuonna 2001. Vuodesta 1991 lähtien tämän työpanoksen kasvu oli ollut runsas 6 prosenttia vuodessa, mikä on jonkin verran korkeampi kasvuvauhti kuin laman pohjalta lähtien saavutettu BKT:n kasvu. Tietosektorilla työskentelevien määrän nostaminen korkeammalle tasolle synnyttää siis endogeenisen kasvun mallissa väliaikaisen kasvun kiihtymisen mutta pysyvän nousun teknologian tasossa. Tämä hyvin pelkistetty endogeenisen kasvun mallikin tuottaa Suomen kannalta tärkeän johtopäätöksen. Suomen innovaatiokyvyn nostaminen lisääntyneillä T&K-panostuksilla on tuottanut viime aikoina teorian lupaamalla tavalla nopeaa kasvua, joka ei jatku, mikäli panostusta ei onnistuta kasvattamaan jatkuvasti. Korkeampi T&K-panoksen taso ei sinänsä ylläpidä tuotannon kasvua, ja tason nostaminen käy ilmeisesti sitä hankalammiksi mitä korkeammalla ollaan. Kasvun haaste on siis oikeastaan vasta edessäpäin.

INNOVAATIOJÄRJESTELMÄN TOIMIVUUS

Kansallisen innovaatiokyvyn tekijöitä ovat yleinen innovaatiotuotannon perusrakenne, yritysklusterien muodostama innovaatioiden soveltamisympäristö ja näiden kahden vuorovaikutus (ks. esim. Furman ym. 2002). Innovaatiokyvyn kannalta on myös syytä erottaa kyky tuottaa uutta tietoa sekä kyky omaksua ja käyttää sitä. Ideasektori ja tuotesektori pitää saada vuorovaikutukseen. Tuotteistaminen ja kaupallistaminen on välttämätöntä taloudellisen menestyksen saavuttamiseksi. Ulkopuolisen rahoituksen lisäystä kaipaavien yliopistojenkin on pitänyt opetella tekemään sitä. Elinkeinoelämän Tutkimuslaitoksen kaltaisia välittäjäorganisaatioita tarvitaan, että tutkimus ja yritystoiminta löytävät paremmin toisensa.

Maailmassa on perustutkimustietoa paljon tarjolla, mutta se on vapaa hyödyke vain niille, jotka ymmärtävät siitä riittävästi. Näin pienenkin maan täytyy pitää huolta riittävästä perustutkimuksesta sekä perustutkintotasoisesta yliopistokoulutuksesta. Vaikka emme olisikaan aivan kärjessä monellakaan alalla, oman omaksumiskyvyn ylläpitämiseksikin tarvitaan jatkuvasti tieteellistä tutkimusta. Samoin tarvitaan tietysti eritasoista koulutusta. Kasvuteorian sanoma menestyksestä ei ole kovin konkreettinen, mutta se selvästikin korostaa inhimillisen pääoman roolia ja samalla todistaa, että talouden institutionaalinen rakenne merkitsee paljon talouden menestykselle.

Tutkimus- ja koulutusinstituutiot ovat tärkeitä, mutta myös yritysten innovaatiohakuisuus merkitsee paljon. Tämä taas syntyy portterilaisen ajattelun mukaisesti siitä, että yritysten toimintaympäristö on kilpailullinen ja sellaisista klustereista muodostuva, että tutkimusjärjestelmän ja yritysten välille voi syntyä hyvät yhteydet. Samalla pääsee syntymään tacit knowledge -tyyppistä tuotannossa opittua inhimillistä pääomaa, joka on vahva valtti kilpailussa. Tämä näyttää olevan hyvin sopusoinnussa myös Suomen kokemusten kanssa. Teleklusteri toimii kansallisen innovaatiojärjestelmän yritysympäristönä, joka pystyy reagoimaan innovaatiokapasiteetin kasvuun. Kokemuksemme on samalla se, että teknologinen osaaminen ei yksin riitä, vaan liiketoimintaosaaminen on välttämätöntä innovaatiotuotannon tehokkuudelle.

Vuoden 1993 nobelisti Douglas C. North on korostanut taloudellisen kasvun perimmäisenä tekijänä innovaatioihin liittyviä omistusoikeuksia. Jos innovaatioista ei seuraa niiden kehittäjälle taloudellista hyötyä, ei innovaatioita myöskään runsain mitoin synny. Tutkijoita kannustaa tieteellinen kunniakin, mutta tuotekehitystä tekevän pitää voida luottaa siihen, että suuret panostukset voidaan saada markkinoilta takaisin. Patentit ja tekijänoikeudet ym. ovat siis tärkeitä asioita. Osaamisstrategian toimivuus taloudellisessa mielessä riippuu institutionaalisesta ympäristöstä. Talouden tuotantokyvyn nousu vaatii tiedon tason nousua sekä sellaisia institutionaalisia olosuhteita, joissa tiedon tason nousu voi realisoitua tuotannon edistykseksi. On siis saatava sekä perustutkimus että T&K-toiminta kehittymään, vaikka perustutkimuksessa tuotetaan julkishyödykkeitä ja tuotekehitystoiminnassa yritysten myymiä markkinahyödykkeitä.

Kuuluuko yliopistossa kehitetty innovaatio yliopistolle vai kehittäjälle, on keskeinen kysymys, josta keskustellaan parhaillaan myös Suomessa. Molempia pitäisi kannustaa innovointiin ja samalla pitäisi huolehtia siitä, että innovaatio selviäisi elävänä markkinoille saakka. Sekä tutkijaa että yliopistoa on siis syytä kannustaa innovaatioiden tuottamiseen. Kriittinen tekijä näyttää olevan se, kumman toimesta paremmin yhteydet yritystoimintaan voidaan järjestää. Tiedepuistoympäristössä nämä yhteydet voidaan järjestää tehokkaasti, vaikka tiedepuistoissa toimivien yritysten T&K-tuotos (patentit ym.) ei välttämättä ole korkeampi kuin niiden ulkopuolella toimivien (Löfsten – Lindelöf 2002, 871).

Tietotuotteiden osalta immateriaaliset omistusoikeudet voivat kuitenkin tehokkuuden kannalta muodostaa myös ongelman. Jos tekijänoikeudet estävät tietotuotteille ominaisen monistettavuuden hyväksikäytön, ne ovatkin taloudellisen tehokkuuden kannalta haitta. Jos tietotuotteisiin sovelletaan kilpailevissa markkinahyödykkeissä tehokkaan ratkaisun tuottavia omistus-

oikeusperiaatteita, voi syntyä monopolin kaltaista tehottomuutta (ks. Romer 2002). Jos taas niille ei turvata ollenkaan omistusoikeuksia, voi syntyä alituotannon takia tehokkuustappiota. Paljon esillä ollut kysymys musiikin vapaasta levityksestä internetissä on hyvä periaatteellinen esimerkkitapaus. Jos tietotuotteisiin sovelletaan tiukkaa tekijänoikeuskontrollia, alitetaan mahdollisuudet hyödyntää uutta teknologiaa. Myöskään julkisen ja yksityisen sektorin työnjako ei ole yksinkertaista näissä asioissa. Julkinen sektori tukee tietotuotantoa laajasti. Tiedon leviämistä ja sen hyväksikäyttöä ei silloin myöskään ole perusteltua rajoittaa tiukasti hinnoittelulla kannustinargumenttiin vedoten.

LIIKETOIMINTAOSAAMINEN VASTAA KASVUHAASTEESEEN

Vuonna 1991 Suomen T&K-menot olivat 2 prosentia BKT:sta ja viime vuonna ne olivat 3,5 prosenttia BKT:sta. Yritysten osuus menoista oli edellisen vuosikymmenen alussa 57 prosenttia ja viime vuonna 71 prosenttia. Tämä merkitsee sitä, että suhteellisesti ottaen Suomi panostaa tutkimukseen nyt enemmän kuin USA tai Japani. Itse asiassa vain Ruotsi on enää meidän edellämme. Suomi voisi hyvinkin olla maailman tutkimusintensiivisin maa aivan näinä aikoina, jos panostus voitaisiin kasvun hiipumisesta huolimatta pitää korkealla, ja käsitys Ruotsin panostuksen vähenemisestä olisi oikea.

Nopean kasvun lisäksi voidaan osoittaa monta seikkaa, jotka puoltavat johtopäätöstä, että osaamisstrategia on tuottanut myös tuloksia. Korkean teknologian vienti on nyt paljon laajempaa kuin tuonti. Patenttitilastoissa olemme nousseet korkealle, ja jopa tieteellisen julkaisutoiminnan määrä ja laatu ovat edistyneet hyvin. Teollisuuden tuottavuus on noussut yli Yhdysvaltojen tason. Suomea ei vaivaa ns. eurooppalainen paradoksi, eli innovaatiokyvyn heikkous suhteessa tutkimuskapasiteettiin (Caracostas – Muldur 1998). Suomi voidaan laajan empiirisen tutkimuksen perusteella mainita Japanin, Saksan ja Ruotsin kanssa eniten innovaatiokapasiteettiaan neljännesvuosisadan aikana nostaneena maana (Furman ym. 2002).

Suomen taloudellista menestystä voi luonnehtia niin, että ns. uuden talouden suunnassa on tapahtunut nopeaa kasvua laman jälkeen. Tälle kasvulle on ollut ominaista ICT-sektorin (information and communication technologies) tuotteiden tuotannon kasvu. ICT:n soveltamisessa talouden muilla sektoreilla ei ole saavutettu merkittäviä tuloksia (ks. esim. Jalava – Pohjola 2002 sekä Koski ym. 2002). Tuottavuuden nousu on ollut nopeaa elektroniikassa muttei muilla aloilla. Mikäli ICT merkitsee laaja-alaista teknisen kehityksen murrosvaihetta, on syytä odottaa, että ICT:n hyväksikäytön puolella

ns. vanhan talouden alueella tapahtuu paljon vielä tulevaisuudessa. Pitää suorastaan toivoa, että Suomen suurista T&K-panostuksista saadaan innovaatiokyvyn nousuna tuottoja vielä tulevaisuudessa. Tässä mielessä uuden talouden loppu ei saisi olla vielä käsillä, vaikka uusi talous -nimityksestä onkin hyvä vähitellen luopua.

Yliopistojen ja koulutuksen rooli Suomen kansallisessa innovaatiojärjestelmässä on ollut tietysti merkittävä. Empiirinen kasvututkimus on osoittanut, että toisen ja kolmannen asteen koulutuksen lisääntyminen nostaa kasvua. Joidenkin tutkimusten mukaan yhden vuoden lisäys keskimääräisessä koulutusajassa lisää kasvua jopa yli prosentilla (Barro 2000, 26). Yliopisto-opetuksen keskimääräisellä pituudella ei kuitenkaan kehittyneiden maiden joukossa ole ollut viime aikoina positiivista vaikutusta tuottavuuden nousuun. Koulutuksen vaikutuksista on vallalla liian yksinkertaisia ja positiivisia näkemyksiä (ks. Krueger – Lindahl 2001). Koulutuksessakin näyttää vallitsevan jonkinlainen vähenevän rajatuottavuuden periaate, ja sen vaikutus yksilötasolla riippuu suhteellisesta asemasta koulutettujen joukossa.

Kehittyneiden maiden innovaatiokapasiteetit näyttävät konvergoituvan kehityksen myötä (Furman ym. 2002, 901). Suomi on hyvä esimerkki maasta, joka on juossut kiinni eturivin etumatkan ja on nyt monen perässä tulevan kiinnostuksen kohteena. Samoin on käynyt kehittyneiden maiden tulotasoille sekä niiden kehitysmaiden tulotasoille, jotka pystyvät panostamaan inhimilliseen pääomaan ja kehittämään institutionaalista rakennettaan (ks. esim. Barro 2000).

Liiketoimintaosaamisen kehittämisestä on alettu puhua vasta viime vuosina, vaikka liikkeenjohdollisen tiedon ja taidon merkitys talouden kehitykselle on tiedetty aina tärkeäksi. Liiketoimintaosaamisen tärkeys nykytilanteessa liittyy teknologisen ja kaupallisen osaamisen vuorovaikutukseen. Innovaatiokapasiteetti muodostuu tutkimusjärjestelmästä ja soveltamisjärjestelmästä, jossa tärkeä tekijä on taloudellisten sovellutusten osalta liikkeenjohdollinen kapasiteetti. Innovaatioinfrastruktuuri johtaa tuloksiin vain sillä ehdolla, että on yritystoimintaa, joka on kiinnostunut ja kykenevä käyttämään hyväksi innovaatiomahdollisuuksia. Tämä ei koske vain tuoteinnovaatioiden tuottamista, vaan myös aloitteellista ja uudistuvaa toimintaa talouden kaikilla sektoreilla. Tutkimus- ja kehitystoimintaedellytysten lisäksi keskeinen edellytys on kilpailullinen toimintaympäristö, joka kannustaa ja suorastaan pakottaa yrityksiä tarttumaan tilaisuuksiin.

Yhdysvalloissa on havaittu – toisin kuin Suomessa – uuden talouden tuottavuusvaikutuksen olleen vahva myös ns. vanhassa taloudessa. Eräs selitys maiden väliseen eroon tarjoutuu luontevasti markkinoiden kilpailulli-

suuden eroista. Suomessa on aloja, jotka eivät ole joutuneet vielä tehokkaasti sisämarkkinakilpailun piiriin. Kaupan piirissä tapahtuneet muutamat markkinoilletulot ovat herättäneet suurta huomiota, mutta ne ovat toivottavasti vasta alkua. Palvelualat toimivat aina paikallisesti ja alueellisesti. Pienet väestökeskittymät eivät tarjoa tehokkuuteen kannustavia toimintaympäristöjä palvelusektorin tuotannolle. Liiketoimintaosaaminenkaan ei yksin riitä, vaan kyse on myös rakenteellisista olosuhteista. Keskeinen tekijä on silloin kilpailua edistävät yhteiskunnan toimet ja kilpailupolitiikan tehokkuus. Kilpailututkimusta pitäisi harjoittaa myös kasvunäkökulmasta (ks. Koponen – Okko –Virtanen 2003).

Usein on epäilty, onko liikkeenjohdon toimintatavoilla lopultakaan suurta vaikutusta yritysten ja talouden menestykseen. Tätä perinteistä kysymystä lähestyttiin selvittämällä, millaisia johtamistekniikoita yrityksissä käytetään ja millaista sijoitetun pääoman tuottoa ja kokonaistuottavuutta yritykset osoittavat. Yhdysvaltoja, Englantia, Saksaa ja Ranskaa koskenut vertailu (Dorgan – Dowdy 2002) osoitti liikkeenjohdollisten käytäntöjen olevan yhteydessä pääoman tuottoasteeseen sekä kokonaistuottavuuteen. Koska voidaan ajatella, että kilpailulliset olosuhteet pakottavat tiettyihin toimintatapoihin, selvitys myös vertaili eri maiden teollisuuden kilpailullisuutta ja liikkeenjohdollisia menetelmiä. Se ei kuitenkaan löytänyt niiden väliltä suoraa yhteyttä, mutta päätyi näkemykseen, että kilpailevissa olosuhteissa hyvät liikkeenjohdon menetelmät johtavat herkemmin hyvään pääoman tuottoon. Vaikka tällainen tarkastelu ei ole riittävä todistamaan näin monisäikeisen ilmiön lopullista luonnetta, se tuo esille kiinnostavan tuloksen ja osoittaa osaltaan, että liiketoimintaosaaminen on keskeinen tuotannontekijä, vaikka tuottavuus toki riippuu muistakin panostekijöistä ja toimintaympäristöstä.

Viime vuosien nopea kasvu syntyi lisääntyvien osaamispanostusten myötä. Se on selitettävissä endogeenisen kasvun idealla. Nostettaessa innovaatiosektorin kapasiteettia kasvu kiihtyi, mutta tämä kiihtyminen ei ole teorian valossa pysyväistä. Vaikka T&K-panostukset pysyisivät uudella korkealla tasollaan, kasvu ei pysy yhtä nopeana, mitä se oli siirtymävaiheen aikana. Normaali suhdannevaihtelu on nyt tekemässä asetelman vaativaksi. Kasvunäkymät ovat hiipuneet, ja tämä osaltaan voi vähentää uskoa osaamisstrategiaan. Liiketoimintaosaamisen kannalta on syytä taas korostaa, että sitä tarvitaan juuri silloin, kun yritystoimintaa hoidetaan vaikeissa oloissa. Nyt kysytään, mitkä ovat kestäviä taloudellisia arvoja ja mitkä ilmiöt edustavat olennaista sisältöä ns. uudessa taloudessa. Näihin kysymyksiin ei saada vastauksia teknologiasta vaan liiketoimintaosaamisesta.

Suomella on taas edessään kasvuhaaste. Suomen menestys on syntynyt taloudellista toimintaa tukevien institutionaalisten rakenteiden avulla. Näitä rakenteitakin on osattava uudistaa globaalia toimintaympäristömme vastaavalla tavalla. Osaamispanostusten pitäminen korkealla tasolla kasvun vaihteluista huolimatta on välttämätön ehto onnistumiselle.

KIRJALLISUUTTA

Aulin-Ahmavaara, Pirkko – Jalava, Jukka (2003): Pääomapanos ja sen tuottavuus vuosina 1975-2001. VATT-Keskustelualoitteita 194, Helsinki.

Barro, Robert J. (2000): Recent developments in endogenous growth theory, The Determinants of Economic Growth, edited by M.S. Oosterbaan, Thijs de Ruyter van Stevennick and N. van der Windt, Kluwer Academic Publishers.

Caracostas, Paraskevas - Muldur, Ugur (1998): Society, The endless frontier, A European Vision of research and innovation policies for the 21st century, European Commission, DG XII, 1998.

Dorgan, Stephen, J. – Dowdy, John (2002): How good management raises productivity, The McKinsey Quarterly, Number 4.

Easterley, William – Levine, Ross (2003): Tropics, germs, and crops: how endowments influence economic development, Journal of Monetary Economics 50, 3-39.

Furman, Jeffrey L. – Porter, Michael E. – Stern, Scott (2002): The determinants of national innovation capacity, Research Policy 31, 899-933.

Giannetti, Mariassunta (2002): The effects on integration on regional disparities: Convergence, divergence or both? European Economic Review 46, 539-567.

Henrekson, M. - Torstensson, J. - Torstensson, R. (1997): Growth effects of European integration, European Economic Review, 41, 1537-1557.

Jalava, Jukka – Pohjola, Matti (2002): Economic growth in the New Economy: evidence from advanced economies, Information Economics and Policy, 14, 189-210.

Jones, Charles I. (1998): Introduction to Economic Growth, New York.

Kangasharju, Aki (1998): Regional differences in Finland: variations in income growth and firm formation, Pellervon taloudellisen tutkimuslaitoksen julkaisuja N:o 17, Helsinki.

Koponen, Aki T. – Okko, Paavo –Virtanen, Martti (2003): Hyödykemarkkinoiden kilpailullisuuden tutkiminen: teoriasta, menetelmistä ja toteuttamisesta, Turun kauppakorkeakoulun julkaisuja, Sarja Keskustelua ja raportteja 1:2003.

Koski, Heli – Rouvinen, Petri – Ylä-Anttila, Pekka (2002): Tieto & talous, Mitä "uudesta taloudesta" jäi, Sitra 253, Edita, Helsinki.

Krueger, Alan – Lindahl, Mikael (2001): Education for growth: Why and for whom?, Journal of Economic Literature 39:4.

Landes, D. (1998): The Wealth and Poverty of Nations, New York.

Löfsten, Hans – Lindelöf, Peter (2002): Science Parks and the growth of new technology-based firms – academic-industry links, innovation and markets, Research Policy, 31, 859-876.

Okko, Paavo (2000): Growth, Human Capital, and Agglomeration Economies, Management Expertise for The New Millennium, edited by Tapio Reponen, Publications of Turku School of Economics and Business Administration, Series A-1:2000. 227-235.

Okko, Paavo (2002): Liiketoimintaosaaminen ja endogeeninen kasvu, Eero Kasanen 50 vuotta, toimittanut Pentti Malaska, Turun kauppakorkeakoulun julkaisuja Sarja C-1:2002.

Pohjola, Matti (1996): Tehoton pääoma, Uusi näkökulma taloutemme ongelmiin, Porvoo.

Puga, Diego (2001): European regional policies in light of recent location theories, Centre for Economic Policy Research, Discussion paper series No. 2767, April 2001.

Ritsilä, Jari (2001): Studies on the Spatial Concentration of Human Capital, Jyväskylä Studies in Business and Economics 7.

Romer, Paul M. (1990): Endogenous Technological Change, Journal of Political Economy, vol. 98, no. 5, pt. 2, S71-S102.

Romer, Paul M. (2002): When Should We Use Property Rights? American Economic Review, May 2002, 213-216.

Routledge, Bryan R. – von Amsberg, Joachim (2003): Social capital and growth, Journal of Monetary Economics 50, 167-193.

Ruuskanen, Petri (2001): Sosiaalinen pääoma – käsitteet, suuntaukset ja mekanismit, VATT-Tutkimuksia 81, Helsinki.

Sala-i-Martin, X. (1994): Regional Cohesion: Evidence and Theories of Regional Growth and Convergence, Centre for Economic Policy Research, London, Discussion paper series No. 1075, November 1994.

Vartia, Pentti – Ylä-Anttila, Pekka (1996): Kansantalous 2021, Elinkeinoelämän Tutkimuslaitos, Sarja B 126.

11. MODELLING FINNISH ECONOMIC GROWTH: 1860-2001*

Timo Teräsvirta** and Dick van Dijk***

Abstract

In this paper we use smooth transition models to describe the dynamic behaviour of annual time series of Finnish GDP, private consumption and investment over the period from 1860 to 2001. The estimated nonlinear models indicate that the Finnish economy enjoyed a stable steady state characterized by high growth rates and that the deep recessions that occurred during this period had exogenous reasons. Furthermore, the recovery from recessions was rapid and endogenous to the dynamic structure of the economy. Not unexpectedly, the models fail to forecast the start of the Great Depression of the 1990s, which appears to have been a rather exceptional period. From 1993 onwards, however, the estimated models generate reasonably accurate forecasts up to five years ahead and correctly predict the recovery from this deep recession.

INTRODUCTION

One of the characteristic features in Pentti Vartia's scientific work has been the desire to look deep into the future and discuss the prospects of the Finnish economy in the long run. In order to make qualified statements about the future it is necessary to understand the past. In a many-faceted work, Vartia and Ylä-Anttila (1996) consider the main factors of Finnish economic growth from 1860 to 1995 and make projections up until the year 2021. Pentti Vartia has also been interested in contemporary Finnish economic history and has investigated the causes and consequences of what is sometimes called the Great

*Financial support from the Jan Wallander's and Tom Hedelius's Foundation for Social Research, Contract No. J02-35, is gratefully acknowledged. Our sincere thanks go to Rolf Maury for constructing the time series used in this work. Any errors or shortcomings in the article are our own responsibility.

** Department of Economic Statistics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden, email: Timo.Terasvirta@hhs.se (corresponding author)

*** Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, email: djvandijk@few.eur.nl

Finnish Depression in the 1990s. The results of this analysis are documented in Kiander and Vartia (1998), which is the most thorough account of those years of economic crisis written so far.

With this paper we have chosen to honour the achievements of Pentti Vartia in these two areas of research, Finnish economic growth and the Great Depression of the 1990s. First, we shall investigate the dynamics of economic growth in Finland using long annual time series of GDP, private consumption and investment that extend from 1860 until 2001. This will be done by employing univariate nonlinear time series models. In particular, we are interested in what the estimated models have to say about the dynamics of switching from expansions to recessions and what they can tell about the mechanism of moving from recessions back to higher growth rates.

Second, we consider the Great 1990s Depression in more detail and examine how surprising an event it may have been. A univariate time series model can hardly be expected to correctly forecast a start of a recession. In fact, as Kiander and Vartia (1998) discussed, forecasters with a considerably larger information set than the past values of GDP failed to foresee the upcoming economic downturn at the end of the 1980s. Instead, we try to find out when it was possible to more or less accurately predict the end of the Great Depression with these models. Naturally, this is done ex post, but as we shall later explain, our forecasts will technically be out-of-sample ones and as such are more realistic than mere in-sample predictions.

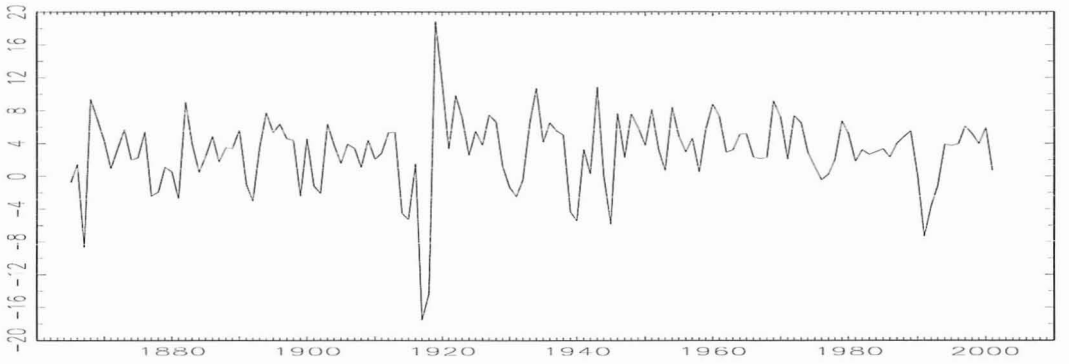
Our analysis bears some resemblance to Skalin and Teräsvirta (1999), who investigated the dynamic behaviour of a number of Swedish macroeconomic variables using long annual time series from 1860 to 1988. As in that work, our statistical tool is the so-called smooth transition autoregressive (STAR) model. This nonlinear model has been found capable of characterizing business cycle asymmetries and other nonlinearities in economic time series. When forecasting with the models, we consider forecast densities as these are more informative than the commonly employed point forecasts. The plan of the paper is as follows. The time series are presented in Section 2. The STAR model and the corresponding modelling cycle are described in Section 3. In Section 4, we

present the estimated STAR models for annual growth rates of GDP, private consumption and investment, and discuss the implied dynamic behaviour of economic growth, in particular the mechanisms of switching between expansions and recessions. In Section 5, we analyze the forecast densities of our three series for the years of the Great Depression, and find that the STAR models adequately forecast the recovery from this deep recession from 1993 onwards. Finally, Section 6 concludes.

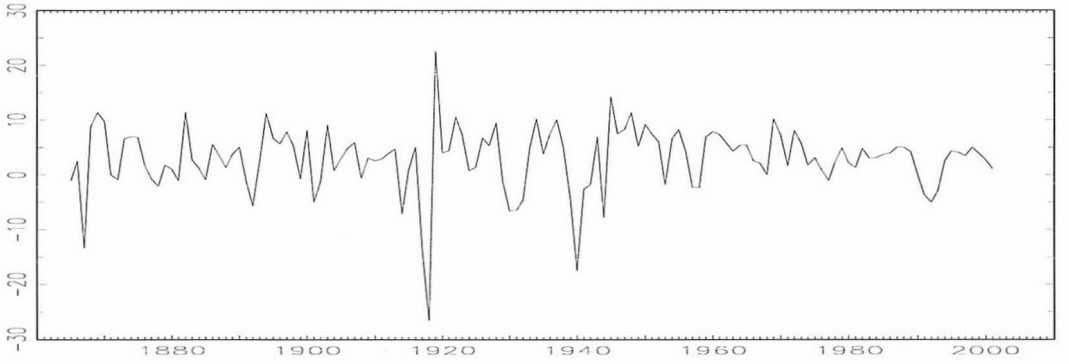
DATA

In this paper we analyze three long annual Finnish time series: the volume index of the Gross Domestic Product, the volume of investments and the volume of private consumption. We have also modelled the volumes of public consumption, exports and imports, but choose not to report the results in this paper. The observations for the years 1860-1960 are based on the figures in Hjerppe (1989), while the ones for the years 1961-1999 have been obtained from the national accounts. The latter contain figures with three different base years; they have been chained together to form a coherent series together with the observations up until 1960. Finally, the observations for the years 2000 and 2001 are preliminary values. For simplicity, they have been given the same weight as the final observations for previous years. A caveat is that in constructing long time series the choice of index may affect the end-product. As we are concerned with first differences of logarithmic volumes, this effect is not as dramatic as it could be if we were dealing with levels. The logarithmic differences have been multiplied by 100 to express the growth rates in percentage points.

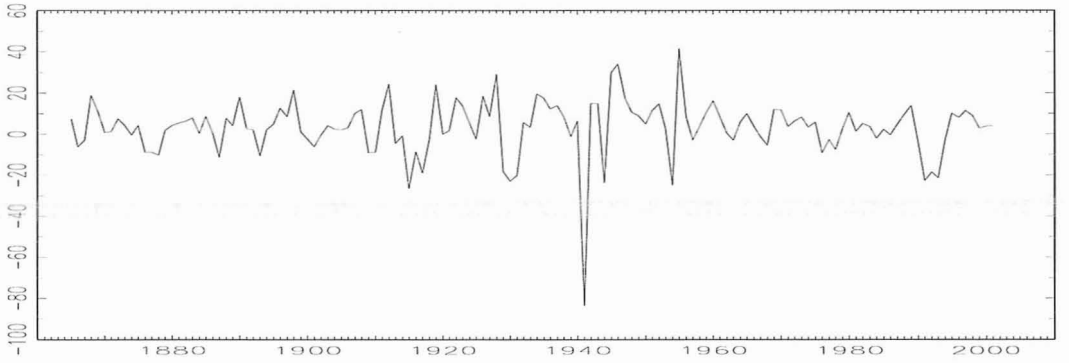
The series contain a few aberrant observations that have been treated separately by introducing dummy variables. For private consumption and GDP, the years 1917 and 1918 are the most obvious outliers: turbulence in Russia in 1917 and the civil war in 1918 following the declaration of independence make these years exceptional. Both variables were also suffering in 1867, the year of starvation in Finland, in 1940, the year of the Winter War, and in 1919, in the aftermath of the two previous exceptional years. In the investment se-



(a) GDP



(b) Private Consumption



(c) Investment

Figure 1: Annual growth rates of Finnish macroeconomic time series.

ries, 1941, marking the start of another war against the Soviet Union, and 1955 are classified as aberrant.

The graphs of the annual growth rates appear in Figure 1. For the GDP and investment series the early 1990s have been a highly unusual period in the Finnish economic history, in particular after excluding 1917, 1918 and 1940. The 1930s recession does not appear as severe by visual comparison. Private consumption is a somewhat different matter: the war years represent the most serious perturbation in that series.

MODELLING WITH STAR MODELS

The statistical tool for our analysis is the smooth transition autoregressive (STAR) model. In the present application, the STAR model has the following form

$$y_t = \boldsymbol{\phi}_1' \mathbf{x}_t + \boldsymbol{\delta}_1' \mathbf{D}_t + \boldsymbol{\phi}_2' \mathbf{x}_t G(s_t; \gamma, c) + \varepsilon_t, \tag{1}$$

where y_t is the annual growth rate of the time series variable of interest, s_t is a stochastic transition variable (see below for discussion of our choice of s_t), $\mathbf{x}_t = (1, \widetilde{\mathbf{x}}_t')'$, $\widetilde{\mathbf{x}}_t = (y_{t-1}, \ldots, y_{t-p})'$, $\mathbf{D}_t$ is a vector of dummy variables for the aberrant observations identified in the previous section, and ε_t is a zero-mean white noise process with variance σ_ε^2. The transition function $G(s_t; \gamma, c)$ is assumed to be given by the logistic function

$$G(s_t; \gamma, c) = (1 + \exp\{-\gamma(s_t - c)/\sigma_{s_t}\})^{-1}, \qquad \gamma > 0, \tag{2}$$

where $\sigma_{s_t} = [\mathrm{var}(s_t)]^{1/2}$ makes γ scale-free. As the value of the transition variable s_t increases, the logistic function changes monotonically from 0 to 1, with the change being symmetric around the location parameter c, as $G(c-z; \gamma, c) = 1 - G(c+z; \gamma, c)$ for all z. The slope parameter γ determines the smoothness of the change in the value of the logistic function. As $\gamma \to \infty$, the logistic function $G(s_t; \gamma, c)$ approaches the indicator function $\mathtt{I}[s_t > c]$ and, consequently, the change of $G(s_t; \gamma, c)$ from 0 to 1 becomes instantaneous at $s_t = c$. When $\gamma \to 0$, $G(s_t; \gamma, c) \to 1/2$ for all values of s_t. The STAR model thus distinguishes two extreme regimes corresponding $G(s_t)$ being equal to 0 or 1. With an appropriate choice for s_t, these regimes may correspond with business cycle recessions

and expansions. For more discussion of the STAR model and its properties, see, for example, Granger and Teräsvirta (1993), Teräsvirta (1994,1998) and Van Dijk, Teräsvirta and Franses (2002).

In applying the STAR model to the present series we follow the modelling cycle consisting of specification, estimation and evaluation, which is briefly outlined in the following. The interested reader is referred to Teräsvirta (1998) or Van Dijk *et al.* (2002) for more detailed expositions. We begin by specifying a linear autoregressive (AR) model for each of the series. Based on the AIC (using a maximum AR order $p = 8$), together with Breusch-Godfrey tests for no error autocorrelation, we arrive at the following orders p of the AR models: $p = 4$ for GDP and investment, and $p = 1$ for private consumption. Next, we test linearity against the STAR model as given in (1) with the transition function now given by

$$G(s_t;\gamma,\mathbf{c}) = (1+\exp\{-\frac{\gamma}{\sigma_{s_t}^k}\prod_{i=1}^{k}(s_t-c_i)\})^{-1}, \qquad \gamma>0, c_1\leq\ldots\leq c_k, \tag{3}$$

using the LM-type tests developed in Luukkonen *et al.* (1988). In this testing approach, the logistic transition function (3) is replaced with a first-order Taylor expansion around the null hypothesis, which can be formulated as H_0 : $\gamma = 0$. This leads to the auxiliary regression model

$$y_t = \boldsymbol{\phi}_1'\mathbf{x}_t + \boldsymbol{\delta}_1'\mathbf{D}_t + \boldsymbol{\beta}_1'\mathbf{x}_t s_t + \boldsymbol{\beta}_2'\mathbf{x}_t s_t^2 + \ldots + \boldsymbol{\beta}_k'\mathbf{x}_t s_t^k + e_t, \tag{4}$$

where the parameter vectors $\boldsymbol{\beta}_j$, $j = 1,2,\ldots,k$, satisfy $\boldsymbol{\beta}_j = \gamma\widetilde{\boldsymbol{\beta}}_j(\boldsymbol{\theta})$, where $\widetilde{\boldsymbol{\beta}}_j(\boldsymbol{\theta})$, $j = 1,2,\ldots,k$, are non-zero functions of the parameters $\boldsymbol{\theta} = (\boldsymbol{\phi}_1', \boldsymbol{\phi}_2')'$. Hence, the original null hypothesis becomes H_0' : $\boldsymbol{\beta}_j = \mathbf{0}$, $j = 1,\ldots,k$, in the transformed equation (4). The standard Lagrange Multiplier [LM] statistic for testing H_0' has an asymptotic χ^2 distribution with $k(p+1)$ degrees of freedom under the null hypothesis. In practice, an F-version of the test is recommended because its size properties in small and moderate samples are much better than those of the χ^2-based test statistic, especially if k and/or p becomes large relative to the sample size. We assume that the transition variable s_t in (3) either is a lagged dependent variable ($s_t = y_{t-d}$, $d > 0$) or, for consumption and investment, a lagged GDP growth rate ($s_t = w_{t-d}$, where w_t is the

GDP growth rate). The reason for considering the lagged GDP growth rate as potential transition variable is that this variable tracks the business cycle perhaps more closely than consumption and investment growth, and we may expect the dynamic properties of the macroeconomic variables under study to be different during expansions and contractions. We test linearity based on (4) for $k = 1, 2$ and 3 and consider values of the delay parameter d in the range $d \in \{1, \ldots, 4\}$. The choice of s_t which gives rise to the smallest p-value of the test for $k = 3$ is taken as the appropriate transition variable. The tests for $k = 1$ and 2 are used to check the appropriateness of the logistic transition function in the models reported below; see Teräsvirta (1994) for details.[1]

In case linearity is rejected, we proceed by estimating logistic STAR models with $k = 1$ in (3). This form of the transition function is most appropriate to distinguish between recessions and expansions, given the choice of transition variables. The estimated models are evaluated by a set of diagnostic checks and misspecification tests developed in Eitrheim and Teräsvirta (1996); see also Teräsvirta (1998).

Finally, it is worth pointing out that in all linearity tests and in the estimated STAR models reported below, the influence of aberrant observations is removed from the transition variables by replacing the growth rates in the relevant years with the mean growth rate. This amounts to assuming that those observations have not been generated by the same process as the remaining ones.

ESTIMATED STAR MODELS

GROSS DOMESTIC PRODUCT

Table 1 contains p-values of the LM-type linearity tests. As can be seen, for the GDP growth rate linearity is rejected at the 5% significance level only when the first lag of GDP is used as transition variable. Setting $s_t = y_{t-1}$ and sequentially removing insignificant coefficients, we arrive at the following spec-

[1]Linearity has also been tested against the alternative of a TV-AR model, which results when $s_t = t$ in (1). Parameter constancy is not rejected at conventional significance levels for any of the series.

ification:

$$y_t = \underset{(0.26)}{-\ 0.63}\ y_{t-2} \underset{(0.061)}{-\ 0.13}\ y_{t-4} \underset{(3.25)}{-\ 12.0}\ D_t^{1917} \underset{(3.27)}{-\ 22.3}\ D_t^{1917} \underset{(3.26)}{-\ 19.4}\ D_t^{1918} \underset{(3.67)}{+\ 10.0}\ D_t^{1919}$$
$$\underset{(0.60)}{(\ 5.13} \underset{(0.29)}{+\ 0.47}\ y_{t-2}) \times G(y_{t-1};\gamma,c) + \widehat{\varepsilon}_t, \tag{5}$$

$$G(y_{t-1};\gamma,c) = (1 + \exp\{\underset{(1.07)}{-\ 2.36}\ (y_{t-1} \underset{(0.84)}{+\ 0.85}\ )/\sigma_{y_{t-1}}\})^{-1}, \tag{6}$$

$\widehat{\sigma}_\varepsilon = 3.20$, $\widehat{\sigma}_{\text{STAR/AR}} = 0.96$, SK $= -0.033(0.44)$, EK $= 0.22(0.30)$, LJB $= 0.30(0.86)$, $\text{LM}_{\text{SC}}(1) = 0.39(0.54)$, $\text{LM}_{\text{SC}}(2) = 0.25(0.78)$, $\text{LM}_{\text{SC}}(4) = 1.18(0.32)$, ARCH(1) $= 3.35(0.067)$, ARCH(2) $= 3.73(0.16)$, ARCH(4) $= 5.66(0.23)$, $\text{AIC}_{\text{STAR/AR}} = -0.044$, $\text{BIC}_{\text{STAR/AR}} = -0.017$,

where OLS standard errors are given in parentheses below the parameter estimates, $\widehat{\varepsilon}_t$ denotes the regression residual at time t, $\widehat{\sigma}_\varepsilon$ is the residual standard deviation, $\widehat{\sigma}_{\text{STAR/AR}}$ is the ratio of the residual standard deviations in the estimated STAR model (5) and the best fitting subset AR model, SK is skewness, EK excess kurtosis, LJB the Lomnicki-Jarque-Bera test of normality of the residuals, $\text{LM}_{\text{SC}}(j)$ is the LM test for no residual autocorrelation up to and including lag j, ARCH(q) is the LM test of no ARCH effects up to and including order q. The numbers in parentheses following the test statistics are p-values. Finally, $\text{AIC}_{\text{STAR/AR}}$ and $\text{BIC}_{\text{STAR/AR}}$ are differences between the Akaike and Schwarz Information Criteria, respectively, of the estimated STAR and the AR models. Negative values indicate that the STAR model is preferred over the AR model.

Further misspecification tests of parameter constancy and no remaining nonlinearity, both in the conditional mean and the residual variance, are reported in Table 2. Figure 2 shows the observed values along with the predictions of the STAR model and the best fitting subset AR model (after removing the identified outliers, such that the differences in fit of the linear and nonlinear models are seen more clearly), the value of the transition function over time, a scatter plot of the transition function values against the transition variable (where each dot corresponds with a single observation), and the sliced spectrum of the STAR model (5) as discussed in Skalin and Teräsvirta (1999).

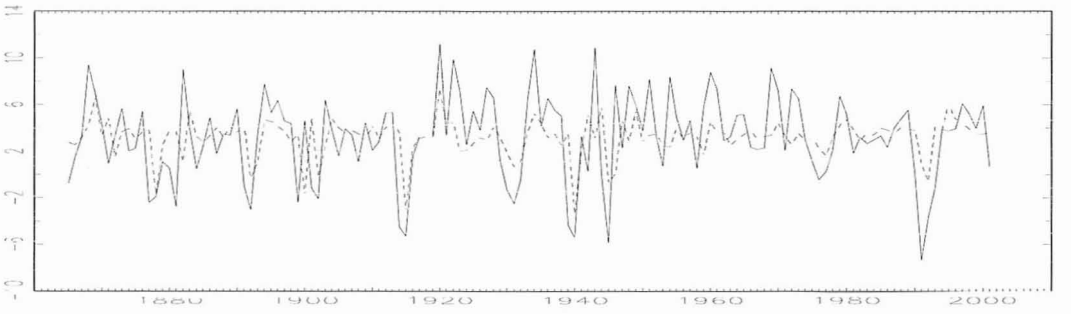
Table 1: Testing linearity of annual growth rates of Finnish GDP, private consumption, and investment

Parameters	$s_t = y_{t-d}$			$s_t = w_{t-d}$		
	LM_1	LM_2	LM_3	LM_1	LM_2	LM_3
GDP ($\widehat{p} = 4$)						
$d = 1$	2.1E−3	0.019	0.028	–	–	–
$d = 2$	0.11	0.10	0.28	–	–	–
$d = 3$	0.18	0.47	0.20	–	–	–
$d = 4$	0.29	0.15	0.19	–	–	–
Private Consumption ($\widehat{p} = 1$)						
$d = 1$	0.48	0.66	0.79	0.69	0.24	0.17
$d = 2$	0.62	0.55	0.71	0.53	0.67	0.029
$d = 3$	0.21	0.16	0.36	0.27	0.22	0.32
$d = 4$	0.059	3.9E−3	4.1E−3	0.11	0.34	0.41
Investment ($\widehat{p} = 4$)						
$d = 1$	3.2E−4	8.5E−4	1.7E−4	6.7E−4	7.6E−5	5.1E−5
$d = 2$	0.23	9.0E−3	1.1E−3	0.094	9.8E−4	5.4E−4
$d = 3$	0.80	0.53	0.16	4.5E−4	9.2E−3	0.011
$d = 4$	0.060	0.097	0.21	4.8E−3	2.4E−3	5.9E−3

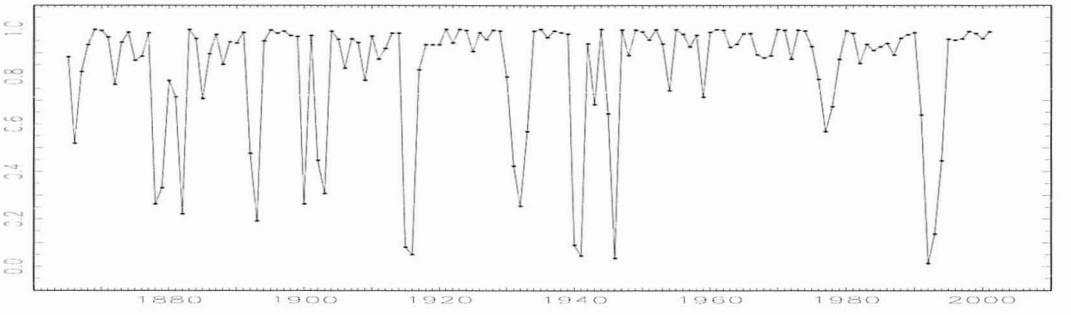
Notes: The table contains p-values of F-variants of the LM_k, $k = 1, 2, 3$, tests of linearity and parameter constancy for annual growth rates of Finnish GDP, private consumption, and investment over the period 1860-2001. Columns headed $s_t = y_{t-d}$ and $s_t = w_{t-d}$ assume that the transition variable in the alternative STAR model is a lagged dependent variable and a lagged GDP growth rate, respectively.

The spectrum is computed only for values of the transition function such that the modulus of the dominant root of the implied AR model is less than 0.98.

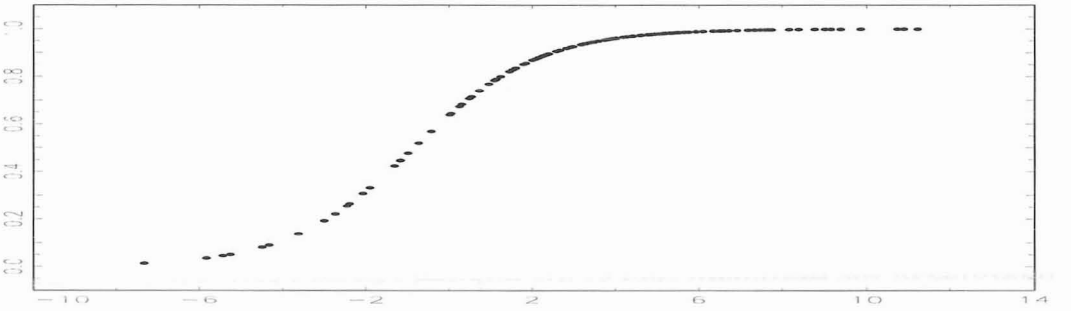
From (5) and (6) it can be seen that growth is quite steady at a high rate and that there does not exist any inbuilt mechanism that would push the economy into a recession. Only an unusually large negative shock or a series of shocks can do that. Recessions have a tendency to be deep, and the recovery occurs through the dynamic structure of the model. There is thus an asymmetry: recessions are caused by shocks exogenous to the model, whereas recovery is due to the dynamics of the process. It is noteworthy that Potter (1995) obtained a similar characterization for US GNP based on threshold autoregressive models. The asymmetry can also be seen from the sliced spectra of equation (5) shown in Figure 2(d). These, however, do not convey the role of the intercepts of equation (5) in the dynamic behaviour of the process. Also note that the



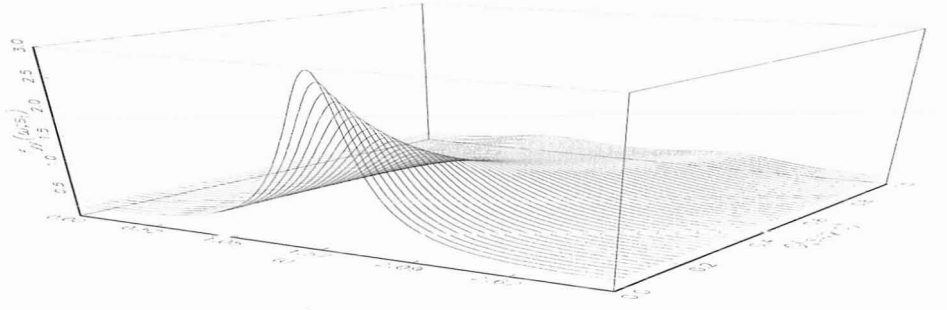
(a) Observed values (solid line), and predictions from STAR model (5) (dashed line) and AR model (dotted line)



(b) Transition function (6) over time



(c) Transition function (6) vs. transition variable Δy_{t-1}



(d) Sliced spectra

Figure 2: Characteristics of STAR model for annual growth rates of GDP as given in (5)-(6).

Table 2: Diagnostic tests of parameter constancy and no remaining nonlinearity in estimated STAR models

Transition variable	$\mathbf{x}_t$			σ_ε^2		
	LM_1	LM_2	LM_3	LM_1	LM_2	LM_3
GDP						
t	0.54	0.66	0.76	0.77	0.44	0.60
y_{t-1}	0.18	0.20	0.10	0.20	0.30	0.49
y_{t-2}	0.28	0.075	0.10	0.84	0.96	0.99
y_{t-3}	0.056	0.21	0.21	0.94	0.22	0.22
y_{t-4}	0.86	0.33	0.56	0.32	0.49	0.53
Private Consumption						
t	0.15	0.31	0.45	0.41	0.088	0.14
y_{t-1}	0.58	0.090	0.19	0.46	0.62	0.49
y_{t-2}	0.84	0.52	0.63	0.82	0.79	0.86
y_{t-3}	0.40	0.42	0.62	0.75	0.93	0.79
y_{t-4}	0.20	0.44	0.51	0.85	0.41	0.54
Investment						
t	0.60	0.65	0.77	0.99	0.19	0.18
w_{t-1}	0.049	0.058	0.14	0.94	0.98	0.91
w_{t-2}	0.97	0.74	0.056	0.36	0.24	0.21
w_{t-3}	0.59	0.36	0.69	0.60	0.80	0.93
w_{t-4}	0.10	0.26	0.15	0.53	0.77	0.82

Notes: The table contains p-values of F-variants of LM diagnostic tests of parameter constancy (rows labelled t) and no remaining nonlinearity (rows labelled y_{t-l} and w_{t-l}, with $l = 1, \ldots, 4$) of intercepts and autoregressive parameters (columns headed $\mathbf{x}_t$) and residual variance (columns headed σ_ε^2) in estimated STAR models for annual growth rates for the Finnish macroeconomic variables.

transition from the extreme "recession regime" ($G = 0$) to the other extreme ($G = 1$) is smooth. This is clear from 2(c) showing the estimated transition function as a function of its argument y_{t-1}.

It can be seen from Figure 2(a), however, that the STAR model does not fully explain the depth of the recession of the 1990s. Its advantage over the linear AR model is mainly due to 1930s and 1940s. From the 1950 onwards, both models fit equally well. The reduction in the residual standard deviation compared to the linear AR model is not higher than 4%. The misspecification tests in Table 2 indicate that the STAR model is adequate, at least in the sense that the hypotheses of parameter constancy and no remaining nonlinearity are not rejected.

PRIVATE CONSUMPTION

Linearity of the annual growth rates in private consumption is rejected with the LM_3 statistic at conventional significance levels for $s_t = y_{t-4}$ and w_{t-2}; see Table 1. The p-value of the test is much smaller when the lagged consumption growth rate is used, while the other linearity tests confirm that this is a more appropriate transition variable than lagged GDP growth. Recursively eliminating insignificant regressors, we obtain the specification:

$$\begin{aligned} y_t = & \underset{(1.13)}{6.44} - \underset{(3.99)}{16.0}\, D_t^{1867} - \underset{(4.00)}{16.5}\, D_t^{1917} - \underset{(4.13)}{33.1}\, D_t^{1918} + \underset{(4.62)}{26.5}\, D_t^{1919} - \underset{(4.04)}{18.5}\, D_t^{1940} \\ & + (\underset{(1.23)}{-4.39} + \underset{(0.074)}{0.23}\, y_{t-1}) \times G(y_{t-4};\gamma,c) + \widehat{\varepsilon}_t, \end{aligned} \tag{7}$$

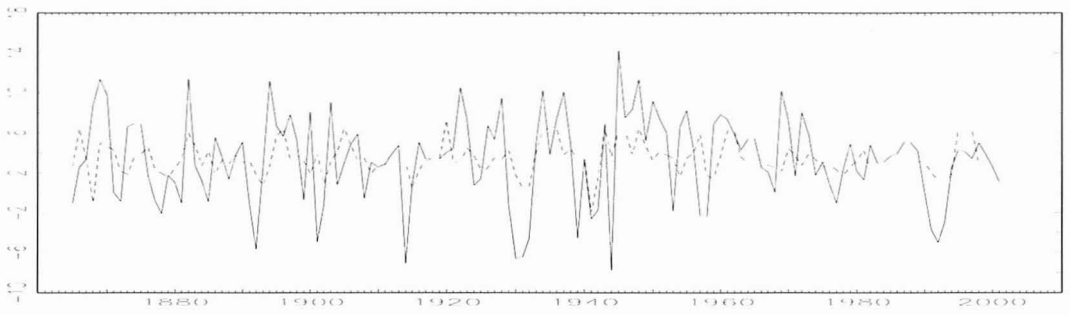
$$G(y_{t-4};\gamma,c) = (1+\exp\{-\underset{(17.3)}{9.86}\,(y_{t-4} + \underset{(0.71)}{1.13})/\sigma_{y_{t-4}}\})^{-1}, \tag{8}$$

$\widehat{\sigma}_\varepsilon = 3.97$, $\widehat{\sigma}_{\text{STAR/AR}} = 0.95$, SK $= -0.18(0.20)$, EK $= 0.079(0.43)$, LJB $= 0.77(0.68)$, $\text{LM}_{\text{SC}}(1) = 2.43(0.12)$, $\text{LM}_{\text{SC}}(2) = 2.71(0.071)$, $\text{LM}_{\text{SC}}(4) = 1.30(0.27)$, ARCH(1) $= 1.01(0.32)$, ARCH(2) $= 1.44(0.49)$, ARCH(4) $= 16.4(2.6\text{E}-4)$, $\text{AIC}_{\text{STAR/AR}} = -0.053$, $\text{BIC}_{\text{STAR/AR}} = 0.011$.

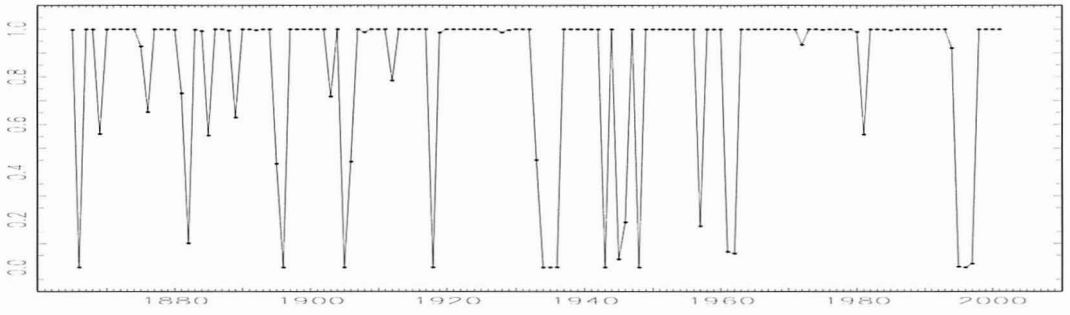
The dynamic structure of the model is quite simple: high growth rates of real consumption are the rule, and there is no built-in dynamic structure slowing down this growth. Consumption also has a tendency to bounce back from recessions quite forcefully: note the high positive value of the intercept when the transition function obtains a value of zero. Nevertheless, the 1990s are not explained in a satisfactory fashion in this case either, see also Figure 3, which indicates that the Great Depression does not fit the previous patterns particularly well. The reduction in the residual standard deviation compared to the linear AR model is 5%.

INVESTMENT

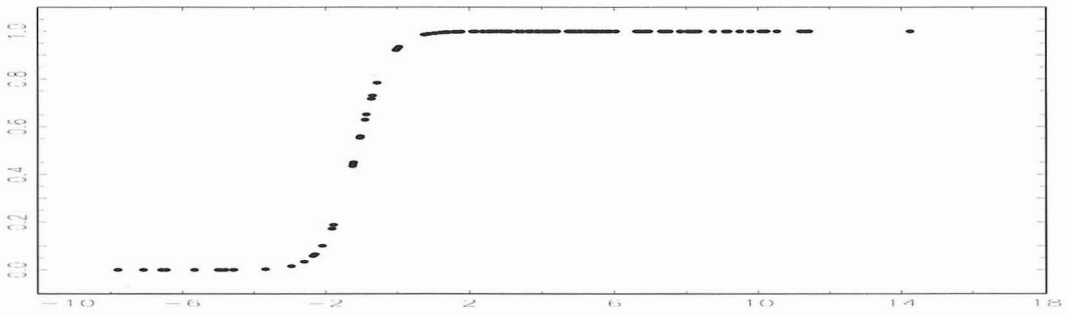
The p-values of the linearity tests in Table 1 show that there is considerable evidence of nonlinear dynamics in investment growth, both when lagged investment or lagged GDP growth are used as transition variable. The strongest



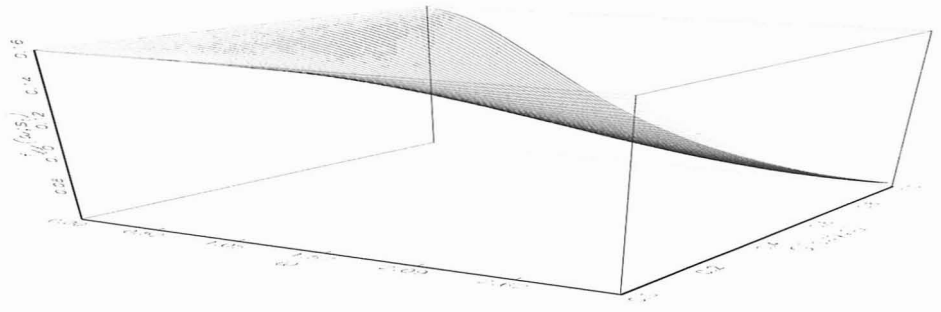
(a) Observed values (solid line), and predictions from STAR model (7) (dashed line) and AR model (dotted line)



(b) Transition function (8) over time



(c) Transition function (8) vs. transition variable Δy_{t-1}



(d) Sliced spectra

Figure 3: Characteristics of STAR model for annual growth rates of private consumption as given in (7)-(8).

rejection of linearity is obtained when the first-lag of GDP growth is considered, and therefore we set $s_t = w_{t-1}$. The final model specification is given by:

$$\begin{aligned} y_t = & \underset{(0.88)}{5.22} + \underset{(0.18)}{0.99}\, y_{t-1} - \underset{(0.22)}{0.75}\, y_{t-3} - \underset{(9.62)}{88.8}\, D_t^{1941} + \underset{(9.47)}{38.1}\, D_t^{1955} \\ & + (- \underset{(0.18)}{0.99}\, y_{t-1} - \underset{(0.061)}{0.11}\, y_{t-2} + \underset{(0.22)}{0.75}\, y_{t-3} - \underset{(0.062)}{0.13}\, y_{t-4}) \times G(w_{t-1};\gamma,c) + \widehat{\varepsilon}_t, \end{aligned} \tag{9}$$

$$G(w_{t-1};\gamma,c) = (1 + \exp\{- \underset{(99.1)}{21.1}\, (w_{t-1} + \underset{(0.34)}{1.24})/\sigma_{w_{t-1}}\})^{-1}, \tag{10}$$

$\widehat{\sigma}_\varepsilon = 9.41$, $\widehat{\sigma}_{\text{STAR/AR}} = 0.90$, SK $= -0.46(0.015)$, EK $= 1.00(8.4\text{E}-3)$, LJB $= 10.5(5.3\text{E}{-3})$, $\text{LM}_{\text{SC}}(1) = 0.52(0.47)$, $\text{LM}_{\text{SC}}(2) = 1.43(0.24)$, $\text{LM}_{\text{SC}}(4) = 1.63(0.17)$, ARCH(1) $= 2.54(0.11)$, ARCH(2) $= 6.85(0.033)$, ARCH(4) $= 8.36(0.079)$, $\text{AIC}_{\text{STAR/AR}} = -0.142$, $\text{BIC}_{\text{STAR/AR}} = -0.057$.

Equations (9) and (10) do not define a pure STAR model as the transition variable is the first lag of GDP growth. The dynamic structure of the model is, however, somewhat similar to that of the STAR model for GDP. For high values of the transition variable (such that $G(w_{t-1},\gamma,c) = 1$), the process is close to being locally white noise. For low values, its most distinct local feature is a pair of complex roots very close to the unit circle with period of about 10 years. This suggests a dynamic structure similar to the one of GDP. There is nothing inherent in the process that would lower the growth rate of investment to recession levels. On the other hand, when investment is growing slowly, there is a tendency for recovery built into the system through the complex pair of roots. This is visible in the sliced spectra in Figure 4(d). Note, however, that a recovery is actually controlled by the transition variable, lagged GDP growth, which in (9) and (10) is strongly exogenous for the parameters of the model.

Common features

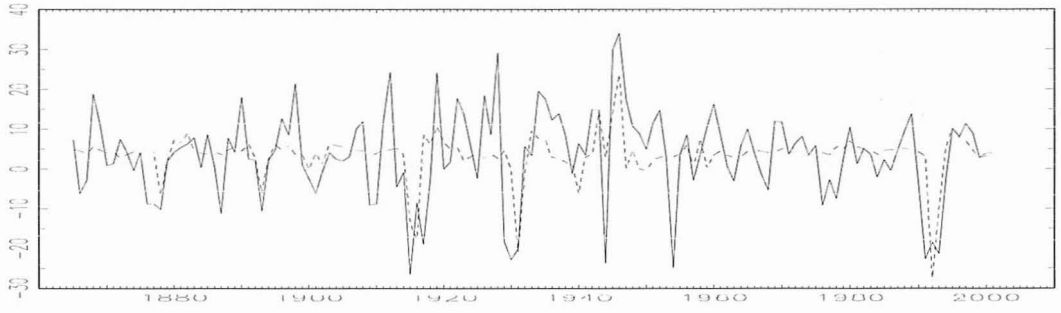
A common characteristic for all three models is that the relatively high growth rates seem to be a steady state that is occasionally disrupted by recessions and crises unleashed by exogenous events. The models suggest that there are no regular fluctuations to speak of at business cycle frequencies in these three variables. It is possible, however, that such fluctuations have existed

for shorter periods of time, but in a model with constant parameters estimated from 140 years of annual data they are overwhelmed by more dramatic events. The dynamic structures of the models also suggest a strong recovery from recessions that hit the Finnish economy. A detailed analysis of the causes of this type of dynamic behaviour is possible only with multivariate models.

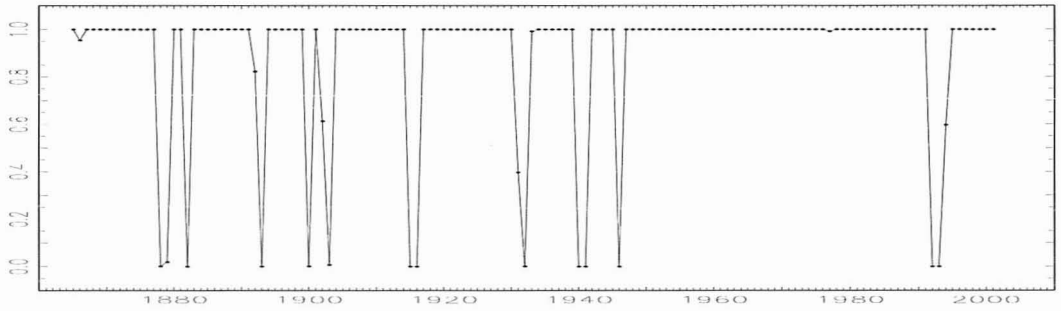
FORECASTING THE RECESSION OF THE EARLY 1990S

In this section we focus on the forecastability of the Great Depression of the 1990s. We only use our estimated STAR models and do not exclude the possibility that someone actually would be able to build a model that would correctly (at least ex post) predict the growth rates we are interested in. The forecasts are generated as follows. The models are re-estimated using observations up to the year at the end of which the forecasts are generated. The models are not respecified, however, which makes the exercise slightly less realistic, but the forecasts are still out-of-sample ones, given the specification. We generate density forecasts up to five years ahead by simulation as discussed in Lundbergh and Teräsvirta (2002), using $N = 10000$ simulated future paths of the variable under consideration. The first base-year is 1989, and the forecasts are then made every year until 1996. Note that in generating forecasts for investment, forecasts are required for the transition variable, the lagged GDP growth rate, for which we take the mean of the density forecasts of this variable.

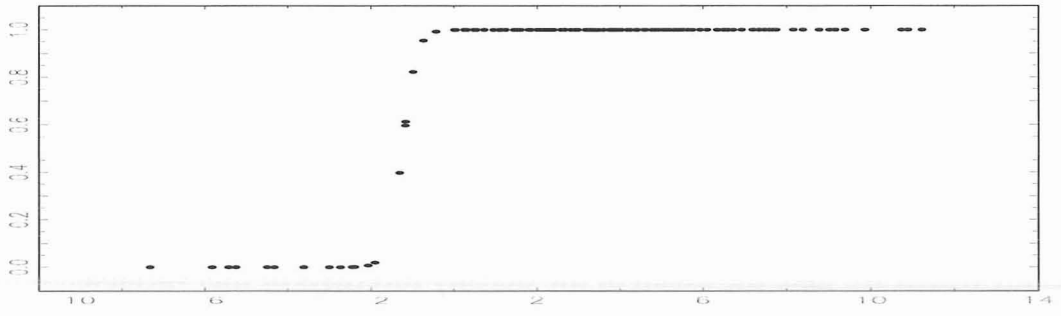
The density forecasts are presented as fan charts following the suggestion of Wallis (1999). They appear in Figures 5-7. Consider first the forecasts for GDP growth in Figure 5. It is clear that the recession comes as a surprise to the STAR forecaster. The true values for 1990-1993 are outside the 90% interval forecasts made in 1989. The situation in 1990 is not any better. The speed of recovery is badly overestimated in 1991 and 1992. The forecasts get back on track in 1993, after which they are slightly pessimistic, but the realized values stay closer to the centre (median) forecast. Even if the model is nonlinear, the forecast densities do not appear asymmetric.



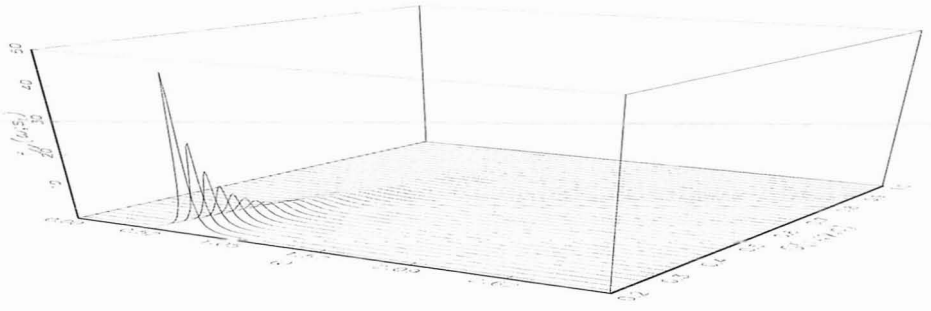
(a) Observed values (solid line), and predictions from STAR model (9) (dashed line) and AR model (dotted line)



(b) Transition function (10) over time

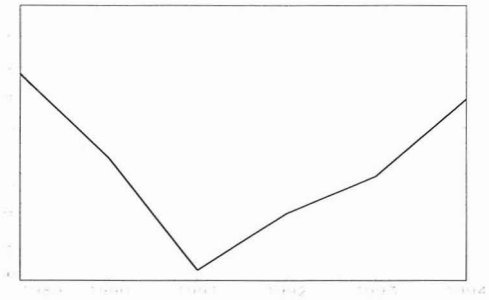


(c) Transition function (10) vs. transition variable Δy_{t-1}

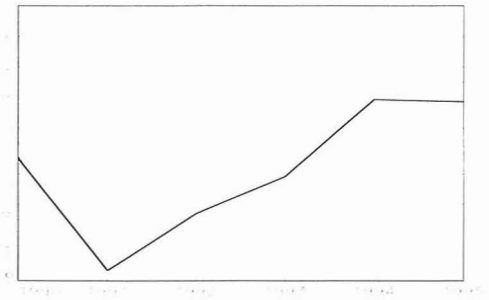


(d) Sliced spectra

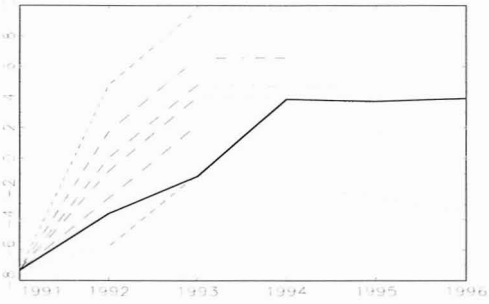
Figure 4: Characteristics of STAR model for annual growth rates of investment as given in (9)-(10).



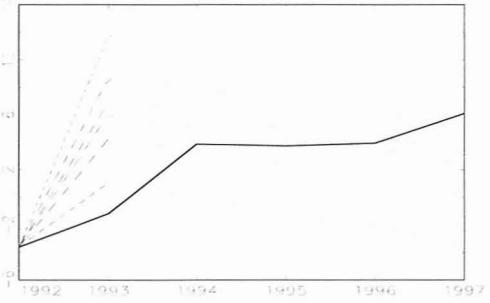
(a) T=1989



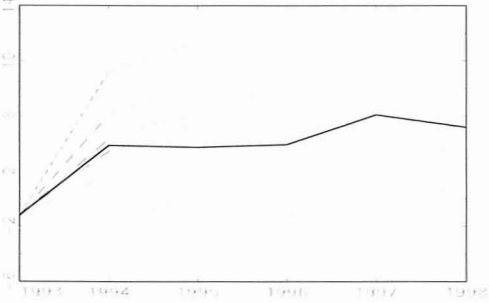
(b) T=1990



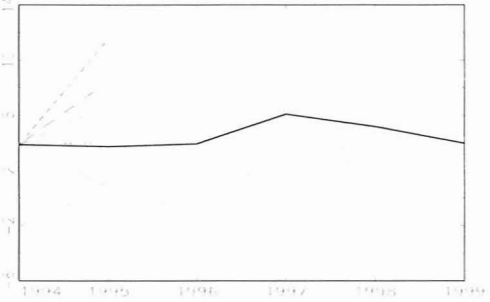
(c) T=1991



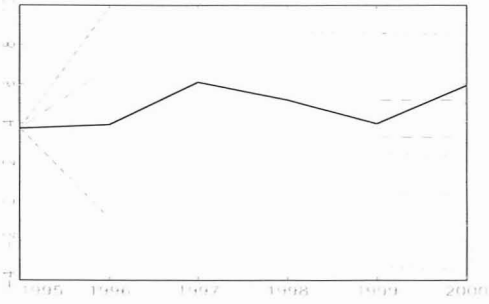
(d) T=1992



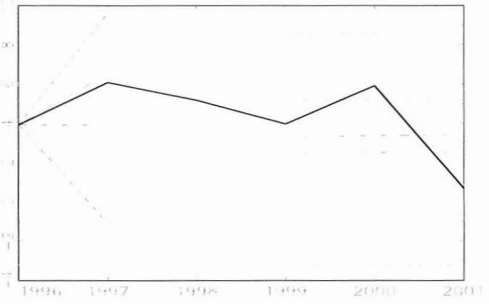
(e) T=1993



(f) T=1994

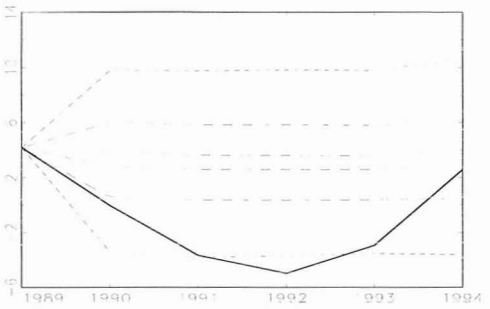


(g) T=1995

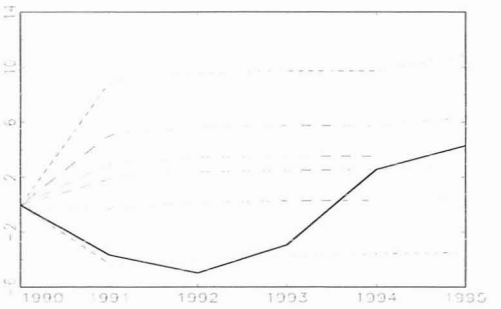


(h) T=1996

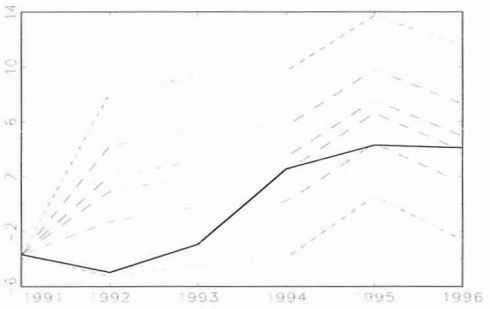
Figure 5: Fan charts for 1- to 5-years ahead forecasts of annual GDP growth rates made at time T, based on the STAR model given in (5) and (6). Long-dashed (sparse-dotted; dot-dashed; dense-dotted; short-dashed) lines indicate the 10% (30%; 50%; 70%; 90%) equal-tail interval. The solid line represents the observed growth rates.



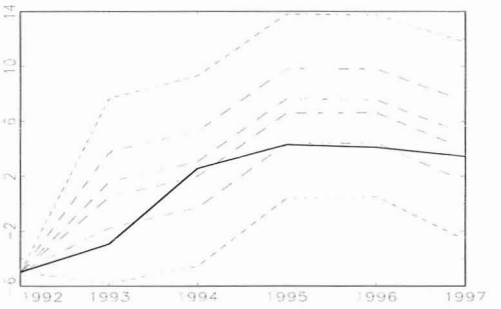
(a) T=1989



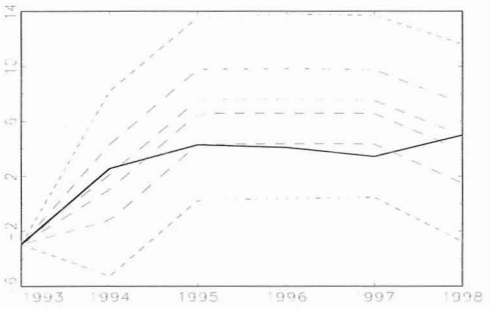
(b) T=1990



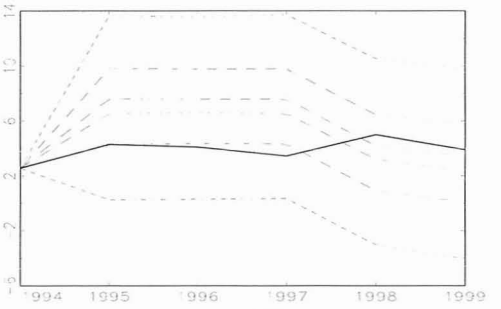
(c) T=1991



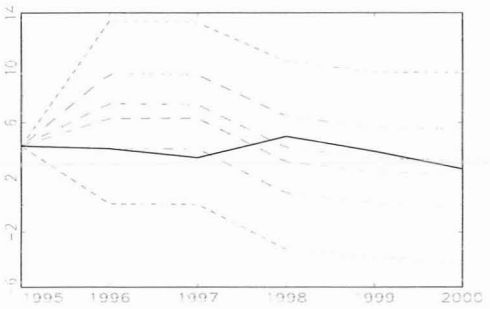
(d) T=1992



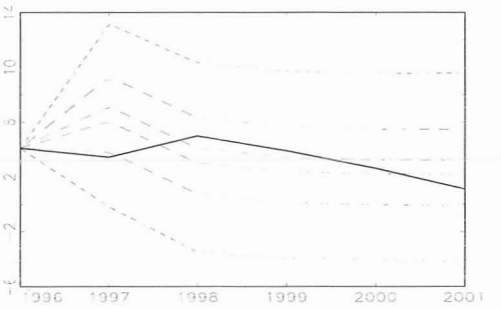
(e) T=1993



(f) T=1994

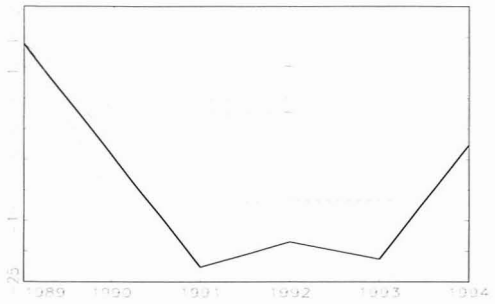


(g) T=1995

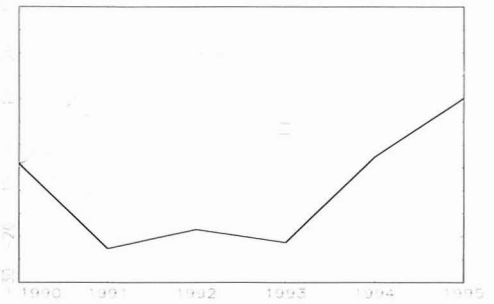


(h) T=1996

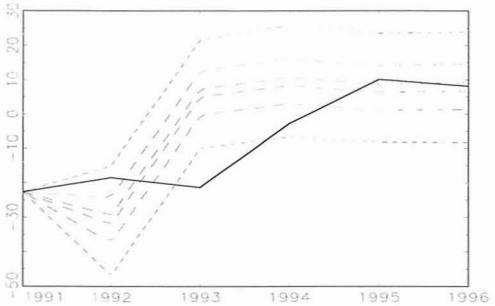
Figure 6: Fan charts for 1- to 5-years ahead forecasts of annual growth rates of private consumption made at time T, based on the STAR model given in (7) and (8). Long-dashed (sparse-dotted; dot-dashed; dense-dotted; short-dashed) lines indicate the 10% (30%; 50%; 70%; 90%) equal-tail interval. The solid line represents the observed growth rates.



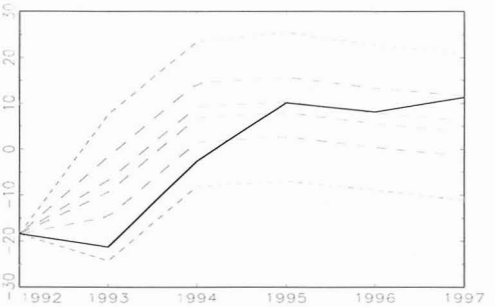
(a) T=1989



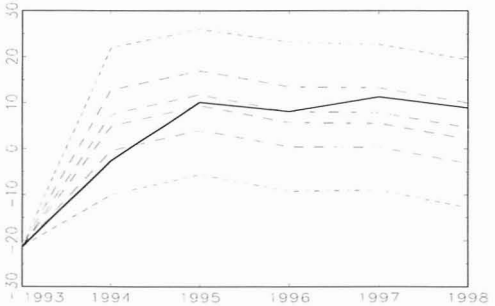
(b) T=1990



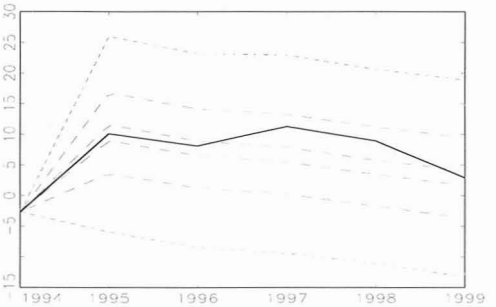
(c) T=1991



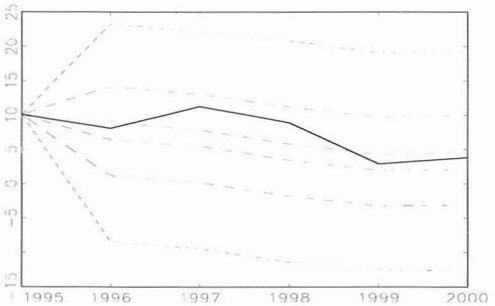
(d) T=1992



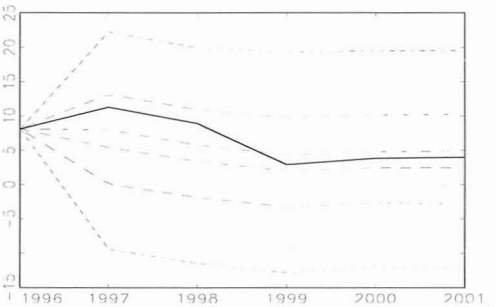
(e) T=1993



(f) T=1994



(g) T=1995



(h) T=1996

Figure 7: Fan charts for 1- to 5-years ahead forecasts of annual growth rates of investment made at time T, based on the STAR model given in (9) and (10). Long-dashed (sparse-dotted; dot-dashed; dense-dotted; short-dashed) lines indicate the 10% (30%; 50%; 70%; 90%) equal-tail interval. The solid line represents the observed growth rates.

Almost the same pattern is found for forecasts of the growth rate of private consumption; see Figure 6. The depth of the slowdown in growth was forecastable neither in 1989 nor 1990, and a recovery was erroneously predicted in 1991. Five-year forecasts from those years are an exception: the short-term drama is being missed but in the long-run the 2% growth rate is about correct. Unlike GDP growth, however, the recovery when it finally arrived was underpredicted in 1993, 1994 and 1995.

As to private investment, the slowdown in investments is correctly predicted in 1989, but the depth of the trough is missed. The forecasts made in 1990 are a complete failure. Even here, the recovery and the stabilization of the growth rate are predicted reasonably well from 1993 onwards. Note that the forecast densities of investment growth are very flat, in other words, the tolerance intervals are remarkably wide.

In general, the results for the early 1990s are not surprising. If economic research institutes with large information sets were not able to see the depression coming, one should not be able to do that with a univariate model. It is noteworthy, however, that the observed values fall even outside the 90% tolerance intervals of the model forecasts as Figures 5-7 indicate.

CONCLUSIONS

In this paper we have modelled three long Finnish macroeconomic time series that have to do with economic growth, the volumes of GDP, private consumption and investment. The estimated nonlinear models indicate that the growth state of the Finnish economy has been reasonably stable and that deep recessions have exogenous reasons. On the other hand, the rapid recovery from recessions is characterized by the dynamic structure of the models.

The models are based on long time series beginning in 1860. Even in this long perspective, the Great Depression of the 1990s appears to be a rather exceptional period. The fit of the models for that period is not particularly good, in the sense that the fitted values are not sufficiently low. Furthermore, out-of-sample forecasts fail until 1993. From that year on, the estimated models generate reasonably accurate forecasts up to five years ahead.

Since the estimated models are univariate, it must be possible to improve upon forecasts from them by constructing multivariate models. Doing that, however, remains outside the scope of the present work.

References

Eitrheim, Ø. and Teräsvirta, T., 1996, Testing the adequacy of smooth transition autoregressive models, *Journal of Econometrics*, 74, 59–76.

Granger, C.W.J. and Teräsvirta, T., 1993, *Modelling Nonlinear Economic Relationships*, Oxford: Oxford University Press.

Hjerppe, R., 1989, *The Finnish economy 1860-1985. Growth and structural change*, Helsinki: Bank of Finland.

Kiander, J. and Vartia, P., 1998, *Suuri lama. Suomen 1990-luvun kriisi ja talouspoliittinen keskustelu*, Helsinki: Elinkeinoelämän tutkimuslaitos.

Lundbergh, S. and Teräsvirta, T., 2002, Forecasting with smooth transition autoregressive models, in Clements, M.P. and Hendry, D.F. (eds.), *A companion to economic forecasting*, Oxford: Blackwell, pp. 485–509.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T., 1988, Testing linearity against smooth transition autoregressive models, *Biometrika*, 75, 491–499.

Potter, S.M., 1995, A nonlinear approach to US GNP, *Journal of Applied Econometrics*, 10, 109–125.

Skalin, J., Teräsvirta, T., 1999, Another look at Swedish business cycles, *Journal of Applied Econometrics*, 14, 359–378.

Teräsvirta, T., 1994, Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association*, 89, 208–218.

Teräsvirta, T., 1998, Modelling economic relationships with smooth transition regressions, in Ullah, A. and Giles, D.E.A. (eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, pp. 507–552.

van Dijk, D., Teräsvirta, T. and Franses, P.H., 2002, Smooth transition autoregressive models - a survey of recent developments, *Econometric Reviews*, 21, 1–47.

Vartia, P. and Ylä-Anttila, P., 1996, *Kansantalous 2021*, Helsinki: Elinkeinoelämän tutkimuslaitos.

Wallis, K.F., 1999, Asymmetric density forecasts and the Bank of England's fan chart, *National Institute Economic Review*, 167, 106–112.

OSA 3:
TALOUSPOLITIIKKA JA TALOUDEN ENNUSTAMISEN VAIKEUS

12. TEOLLISUUSTUOTANNON KÄÄNNEPISTEIDEN ENNAKOINTI BAROMETRIAINEISTON AVULLA

Jussi Mustonen ja Penna Urrila*

JOHDANTO

Ajantasaisen talouskehitystä kuvaavan tiedon määrä ja kysyntä on lisääntynyt niin yrityksissä, mediassa kuin talouspolitiikasta vastaavien piirissä viime vuosina. Suhdannekehitystä kuvaavia tietoja saadaan nykyään monista lähteistä, kuten virallisista tilastoista ja kyselytutkimuksiin pohjautuvista suhdanneindikaattoreista.

Tilastoluvuilla, suhdanneindikaattoreilla ja eri ennustelaitosten makroennusteilla on oma osansa suhdannekehitystä kuvattaessa. Ne eivät ole toisiaan poissulkevia vaan pikemminkin täydentäviä. Kyselytutkimuksiin perustuvien suhdanneindikaattoreiden etuina moniin virallisiin tilastolukuihin verrattuna on niiden nopeus sekä ennakoiva luonne. Niillä on myös mahdollista kartoittaa taloudenpitäjien odotuksia.

Makrotalousennusteiden tyypillinen ennuste- ja tarkastelujakso on pitkähkö, yleensä vähintään kalenterivuosi, ja ennusteita tehdään usein 1-3 vuotta eteenpäin. Suhdanneindikaattoreiden ennakointihorisontti ulottuu sen sijaan korkeintaan vain noin 6-9 kuukauden päähän. Ajantasaisen suhdanneseurannan tarpeisiin makroennusteiden tekeminen harvakseltaan, tyypillisesti neljännesvuosittain, ja ennustejaksojen sitominen kalenterivuoteen on usein riittämätöntä. Toisaalta mikään ei estä tiheämpirytmisen suhdanneuran tarkastelua vuosiennusteiden lisäksi – ja näin on myös usein tehtykin – vaikka pääpaino on vuosimuutoksissa.

Ennustevirheiden raportointi antaa kuvan ennusteiden osuvuudesta, ja monissa ennustelujulkaisuissa, kuten ETLAn Suhdanne -julkaisussa, tästä on tullut rutiinia. Vaikka makrotalousennusteiden menetelmät kehittyvät jatkuvasti, niiden osuvuus on kuitenkin ollut vain kohtalaista. Erityisen vaikeaa on yleensä ollut ennakoida taantumien ajoitusta, vrt. esim. Loungani (2000).

* Jussi Mustonen on johtaja ja Penna Urrila talouspoliittinen asiamies Teollisuuden ja Työnantajain Keskusliitossa (TT).

Yritysten strategisen suunnittelun tarpeisiin, samoin kuin talouspoliittisessa päätöksenteossa, on vuosimuutosennusteiden sijaan usein tärkeämpää saada ennakkovaroituksia kasvu- tai taantumajaksojen alkamisista ja päättymisistä. Juuri käännepisteiden ennakoimisessa suhdanneindikaattorit tarjoavat mielestämme mahdollisuuksia, jotka ovat osin vielä hyödyntämättä.

Tässä artikkelissa tarkastelemmekin suhdannekäänteitä ja niiden ennakoimista suhdannekyselyaineistojen avulla. Tavoitteena on kehittää Suomen aineistolla todennäköisyystyyppinen talouden käännepisteitä ennakoiva malli – käännepisteindikaattori – joka soveltuisi mahdollisimman hyvin teollisuustuotannon käännepisteiden ennakointiin.

Artikkelimme rakenne on seuraava: ensin pohdimme suhdannekäänteiden ja käännepisteiden ennakoinnissa käytettyjen menetelmien ja indikaattoreiden määrittelyä sekä käyttöä. Sen jälkeen kuvaamme mielenkiinnon kohteena olevan teollisuustuotannon suhdannesarjat sekä niiden käänteiden ennakointiin käytettävät TT:n suhdanneindikaattorit. Aineiston kuvauksen jälkeen esittelemme valitsemamme ennakointimenetelmän keskeiset piirteet ja mallinvalinnan ja estimoinnin vaiheet. Seuraavaksi pohdimme ja perustelemme käyttämämme lähestymistavan taustalla olevan tappiofunktion ominaisuuksia ja tämän jälkeen esittelemme varsinaiset tulokset. Lopuksi teemme yhteenvedon sekä alustavia johtopäätöksiä ja pohdimme mahdollisia jatkotutkimusten kohteita.

SUHDANNEKÄÄNTEET TALOUDEN SEURANNASSA

Talouden kasvunopeuden mittaamiseksi ei ole varsinkaan neljännesvuosi- tai kuukausitasolla yksikäsitteisiä, yleisiä tapoja. Esimerkiksi Yhdysvalloissa on yleistä tarkastella BKT:n neljännesvuosisarjojejn kausitasoitettuja muutoksia, jotka on korotettu vuositasolle. Euroopassa taas on ollut tapana analysoida tuotannon muutoksia edelliseen vuoteen verrattuna myös neljännesvuosisarjoilla.

Ajantasaisen suhdanneseurannan kannalta peräkkäisten neljännesvuosilukujen muutosten käyttäminen on perusteltua. Vuoden yli tapahtuvissa muutostarkasteluissa suhdannekäänteet siirtyvät myöhäisemmäksi kuin alkuperäisessä aikasarjassa. Ongelmaksi voi neljännesvuosimuutosten käytössä tosin muodostua aikasarjojen voimakas kausivaihtelu. Suomen kaltaisessa pienessä taloudessa aikasarjat heilahtelevat hyvinkin satunnaisesti, eikä kausitasoitusta ole mahdollista tehdä samalla luotettavuudella kuin esim. USA:ssa.

Myös suhdanteiden käännepisteiden määrittäminen on ongelmallista. Ei ole yksiselitteistä vastausta sille, millainen muutos suhdannevaihteluita kuvaavassa aikasarjassa pitäisi tulkita suhdannekäänteeksi. Yhdysvalloissa on NBER:n alainen komitea, joka jälkikäteen määrittää suhdannekäänteiden tapahtumisajankohdan. Komitea seuraa useita aikasarjoja, ja käänteen tulkitaan tapahtuneen vasta kun riittävän monissa sarjoissa on tapahtunut selvä suhdannekäänne, ks. tarkemmin Moore (1983).

Perinteisen suhdannevaihtelun (classical business cycle) rinnalla puhutaan myös kasvusyklistä (growth cycle). Perinteisessä suhdannevaihtelussa taantumaksi luetaan siis ainoastaan tilanteet, joissa tuotanto on absoluuttisesti supistunut, kun taas kasvusykleissä taantumaksi katsotaan tilanne, jossa tuotannon kasvu on jäänyt selvästi pitkäaikaisen trendin alapuolelle.

Varsin yleinen tapa käännepisteiden identifioinnissa on käyttää kahden havainnon nyrkkisääntöä: käänne tapahtuu, kun (taloudellisen neljännesvuosisarjan) kahden peräkkäisen havainnon muutoksen etumerkki vaihtuu ja säilyy sen jälkeen samana myös seuraavassa havainnossa. Tätä yksinkertaista tapaa olemme soveltaneet alla olevassa empiirisessä osassa talouden luokittelussa kasvu- ja taantumajaksoihin.

KÄÄNNEPISTEINDIKAATTOREIDEN KÄYTÖSTÄ

Perinteisesti suhdanneindikaattoreilla on kuvattu suhdannekehitystä yleisellä tasolla. Indikaattoreista on tehty päätelmiä suhdannekierron vaiheesta ja erilaisten yksinkertaisten päätössääntöjen kautta niiden kehityksestä on pyritty myös ennustamaan suhdannekäänteitä.

Todennäköisyystyyppinen suhdanneindikaattori, jonka keskeisenä tehtävänä on ennustaa juuri suhdannekäänteitä, esiteltiin ensimmäistä kertaa artikkelissa Neftçi (1982). Kyse ei itse asiassa ole mistään uudesta indikaattorista, vaan artikkelissa käytettiin ns. jonotodennäköisyysmenetelmää (sequential probability method). Menetelmän avulla muodostettiin perinteisestä yrityskyselydatasta lasketusta suhdanneindikaattorista uusi indikaattori, joka kuvasi todennäköisyyttä, että talous on kasvussa.

Tämä indikaattori oli estimoitu USA:n 1950-1970-lukujen aineistolla, ja sen antamat tulokset olivat varsin lupaavia. Indikaattorin osuvuus riippuu kuitenkin ratkaisevasti taustalla olevan indikaattorin tarkkuudesta. Jonotodennäköisyysmenetelmän etuna on kuitenkin se, että se suodattaa alkuperäisestä indikaattorista äkilliset heilahdukset pois, ja antaa signaalin suhdannekäänteestä vasta, jos riittävän moni peräkkäinen havainto osoittaa suhdan-

nenäkymien selvää muutosta. Suomen aineistolla vastaava indikaattori on esitetty teoksessa Urrila (2001).

Tässä artikkelissa olemme päätyneet rakentamaan ja estimoimaan samankaltaisen todennäköisyystyyppisen indikaattorin kuin Koskisen ja Öllerin Ruotsin aineistoon soveltama käännepisteindikaattori (Koskinen ja Öller, 1998). Koskinen ja Öller kutsuvat suhdannekäänteiden ennakointiin tarkoitettua indikaattoriaan Markov-Bayesilaiscksi luokittelijaksi. Menetelmän yksi etu on, että se mahdollistaa ennakoinnissa käytettävien useampien aikasarjojen sisältämän informaation yhdistämisen helposti.

Ruotsissa Konjunkturinstitutet on jatkanut Koskisen ja Öllerin kehittelemän indikaattorin julkaisemista säännöllisesti. Neljä kertaa vuodessa julkaistavat ajantasaista suhdannetilannetta kuvaava "samtida vändpunktsindikator" ja yhden vuosineljänneksen suhdannekehitystä ennakoiva "framtida vändpunktsindikator" ovat osoittautuneet varsin luotettaviksi suhdannekuvaajiksi myös viime vuosien nopeissa suhdannevaihteluissa (ks. Konjukturinstitutin kotisivu www.konj.se).

TT:N SUHDANNEKYSELYAINEISTO JA KÄYTETYT VERTAILUSARJAT

Ensimmäinen suomalainen suhdannetiedustelu on Suomen Teollisuusliiton vuonna 1966 aloittama Suhdannebarometri. Nykyisin se on osa Teollisuuden ja Työnantajain Keskusliiton TT:n suhdannetiedusteluja. Suhdannebarometrin otokseen kuuluu nykyisin noin 700 teollisuusyritystä, yli 100 rakennusyritystä ja yli 300 palveluyritystä. Kyselyn vastaajayritykset edustavat valtaosaa suomalaisesta elinkeinoelämästä.

Neljä kertaa vuodessa tehtävän Suhdannebarometrin rinnalla aloitettiin vuonna 1993 myös kuukausittainen suhdannetiedustelu. Vuodesta 1995 lähtien TT:n suhdannetiedustelut ovat olleet osa EU:n harmonisoitua suhdannetiedustelujärjestelmää.

Suhdannebarometrin kyselylomakkeessa on noin 40 kysymystä (kuukausitiedustelussa 10), jotka kaikki ovat potentiaalisia suhdanneindikaattoreita. Selvästi yleisimmin suhdanneindikaattorina on käytetty "suhdannenäkymät lähitulevaisuudessa" -kysymystä, jossa vastaajat arvioivat lähitulevaisuuden yleistä suhdannekehitystä yrityksensä kannalta. Suhdannenäkymäindikaattori vastaa esimerkiksi IFO-instituutin Geschäftserwartungen -indikaattoria, joka on toinen komponenttisarja IFO-indeksissä.

Suhdannebarometrin vastausten käsittely on menetelmiltään melko yksinkertaista. Vastaukset painotetaan vastaajien liikevaihdolla (henkilöstöä

koskevissa kysymyksissä henkilöstöllä) ja lasketaan kunkin vastausluokan suhteellinen osuus vastanneista. Suhdannekyselyjen tulokset on lisäksi esitetty ns. saldolukuina. Saldoluku lasketaan vähentämällä positiivista kehitystä kuvanneiden tai ennustaneiden vastaajien prosenttiosuudesta negatiivista kehitystä kuvanneiden prosenttiosuus, joten saldolukuaikasarjat vaihtelevat välillä (-100, 100).

Saldolukujen voi tulkita sellaisenaan kuvaavan muutosta, joten niitä voi tarkastella rinnan esimerkiksi tuotannon muutosta kuvaavien aikasarjojen kanssa. Saldolukujen käyttöä on perusteltu lähinnä niiden selkeydellä, koska niitä käytettäessä kyselyn keskeinen sisältö saadaan tiivistettyä yhteen aikasarjaan. Saldolukuja ei kuitenkaan pidä sekoittaa prosenttimuutoksiin, kuten joskus on virheellisesti tehty: saldoluvut kuvaavat muutosta odottavien vastaajien (netto-)osuutta vastausjoukosta, ei muutoksen voimakkuutta.

Pelkkien saldolukujen käyttö suhdanneindikaattoreina ei ole täysin ongelmatonta, koska osa vastauksiin sisältyvästä informaatiosta jää huomioimatta. Saldoluvun käytön vaikutuksesta suhdanneindikaattoreiden ominaisuuksiin on julkaistu jonkin verran tutkimuksia. Teräsvirta (1985) on valinnut ennusteiden selittäviksi muuttujiksi saldolukujen sijaan yksittäisten vastausvaihtoehtojen suhteelliset osuudet. Bergströmin (1993) mukaan tällä ei pystytä oleellisesti parantamaan mallien tarkkuutta saldolukujen käyttöön verrattuna. Öller (1992) toteaa, että saldolukujen käyttö on perusteltua ainoastaan silloin, jos vastausluokan "ennallaan" suhteellinen osuus ei ole voimakkaasti korreloitunut vastausluokkien "suurempi" ja "pienempi" osuuden kanssa. Hänen mukaansa tämä ei useinkaan toteudu, ja siksi olisi parempi käyttää suhteellisia osuuksia saldolukujen sijaan.

Kangasniemen ja Tsuparin (1997) tutkimuksessa on tarkasteltu TT:n Suhdannebarometrin suhdannenäkymäkysymyksen eri vastausluokkien välistä korrelaatiota sekä saldoluvun toimivuutta suhdanneindikaattorina vastausluokkien suhteellisiin osuuksiin verrattuna. Yhteenvetona tässä tutkimuksessa todettiin, että Suomen aineistolla vastausluokan "ennallaan" suhteellinen osuus korreloi heikosti vastausluokan "suurempi" suhteellisen osuuden kanssa. Sen sijaan ko. vastausluokan suhteellinen osuus korreloi kohtalaisesti luokan "pienempi" suhteellisen osuuden kanssa. Näin ollen Öllerin (1992) tutkimuksessa asetetut ehdot eivät täysin toteudu.

Kangasniemen ja Tsuparin (1997) tutkimuksessa havaitaan kuitenkin, että vastausluokkien suhteelliset osuudet eivät ole suhdanneindikaattoreina juurikaan saldolukuja parempia. Tämän vuoksi olemme jatkossa ensisijaisesti käyttäneet juuri saldolukuja todennäköisyystyyppisten indikaattoreiden etsimisessä.

Suhdannebarometrin käytöstä talouden ennustamisessa on tehty TT:ssä ja myös ETLAssa useita tutkimuksia, ks. Kangasniemi ja Tsupari (1997), Takala ja Tsupari (1999) ja Urrila (1999, 2001).

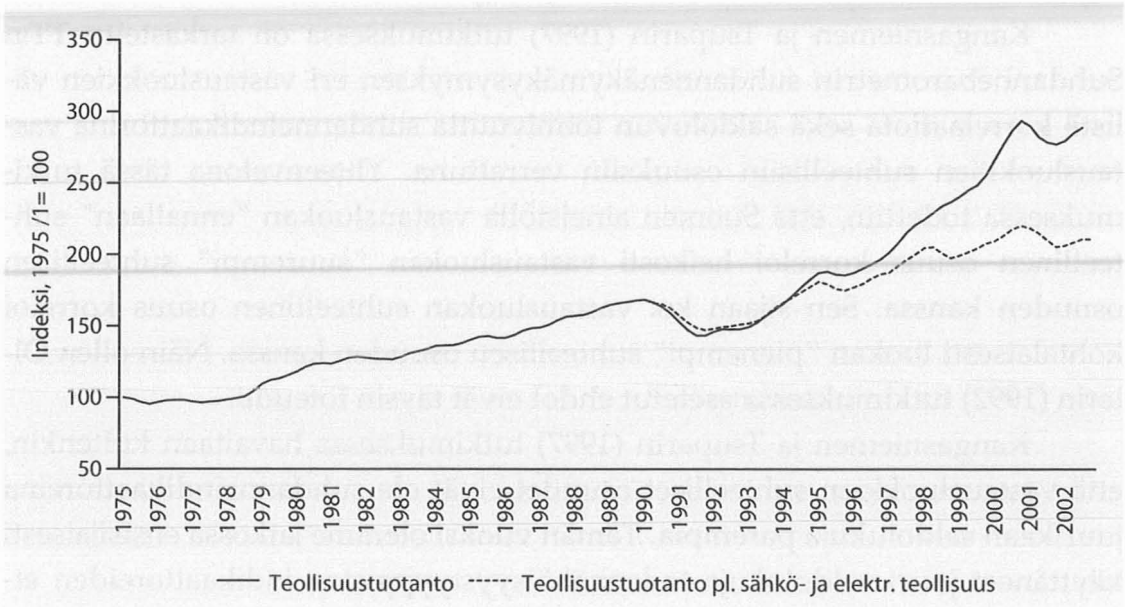
Jatkossa keskitymme Suomen teollisuustuotannon neljännesvuosimuutoksia ennakoivan käännepisteindikaattorin kehittämiseen. Olemme valinneet suhdannevaihteluita kuvaavaksi vertailuaikasarjaksi kansantalouden tilinpidon mukaisen tehdasteollisuuden tuotannon volyymia kuvaavan aikasarjan trendin. Trendi on laskettu alkuperäisestä sarjasta X12-menetelmällä. Trendin neljännesvuosimuutoksien perusteella on määritelty sarjan käännepisteet ja kasvu- sekä taantumajaksot käyttäen kahden peräkkäisen merkin sääntöä.

Käytetty ajanjakso on 1975/Q1-2002/Q3. Olemme säästäneet jakson kahdeksan viimeistä vuotta mallin arviointiin, ja estimointiperiodina on käytetty vuosia 1975-1994.

Vertailun vuoksi olemme myös määrittäneet käännepisteet sellaiselle teollisuustuotantosarjalle, josta on poistettu elektroniikka- ja sähköteollisuus (toimialaluokka DL). Kyseinen aikasarja on saatavilla teollisuustuotannon kausitasoitetusta volyymi-indeksistä vuosille 1990-2002. Tästä aikasarjasta on laskettu trendin muutos neljännesvuosittain edelliseen neljännekseen verrattuna. Sarja on ketjutettu vuosia 1975-1989 kuvaavan Kansantalouden tilinpidon teollisuustuotantosarjan kanssa.

Kuvioon 12.1 on piirretty koko teollisuutta kuvaava sarja (ylempi sar-

Kuvio 12.1. Teollisuustuotanto ml. ja pl. elektroniikka- ja sähkötuotteiden valmistus



ja) sekä sarja, josta on poistettu elektroniikka- ja sähkötuotteiden valmistus (DL). Sarjat erkanevat selvemmin toisistaan 1990-luvun puolivälissä, minkä jälkeen elektroniikka- ja sähköteollisuuden tuotannon muutokset ovat dominoineet aggregaattisarjan kehitystä. On kuitenkin mielenkiintoista arvioida, miten elektroniikka- ja sähköteollisuudesta puhdistetun teollisuuden tuotanto ja suhdanteet ovat kehittyneet ja miten ennakointi-indikaattorit ovat niiden suhteen toimineet.

MARKOV-BAYESILAINEN LUOKITTELIJA

Olemme valinneet ennakointimallimme rakentamiseen todennäköisyystyyppisen suhdanneindikaattorin menetelmän, jonka ovat kehittäneet Koskinen ja Öller (1998). He kutsuvat menetelmää Markov-Bayesilaiseksi luokittelijaksi, ja se perustuu piilotettujen Markov-mallien teoriaan (Hidden Markov Models, HMM). Oma sovelluksemme poikkeaa Koskisen ja Öllerin vastaavasta lähinnä aikasarjojen valinnan ja käyttötarkoituksen suhteen, koska – toisin kuin he – emme etsi vuosimuutosten käänteitä vaan alkuperäisen sarjan trendin käänteitä. Toisaalta sovelluksemme eroaa jonkin verran mallinvalinnassa ja estimoinnissa käytettävän tappiofunktion käsittelyn osalta.

Oheinen kuvaus itse menetelmästä noudattaa pääosin artikkelissa Koskinen ja Öller (1998) esitettyä merkintätapaa:

Tarkastellaan kahta erilaista maailmantilaa kuvaavaa luokkaa, $i \in \{1,2\}$ ja havaintovektoria y, joka halutaan luokitella kuuluvaksi jompaankumpaan luokista. Olkoon pari $Z = (Y, J)$, jossa $Y \in R^n$ on satunnaisvektori ja $J \in \{1,2\}$ on satunnaismuuttuja, joka liittää luokitteluinformaatiota Y:hyn. Vain Y voidaan havaita, kun taas J ei ole havaittavissa. Tällöin tarvitaan luokittelusääntö (funktio) $g : R^n \rightarrow \{1,2\}$, joka mahdollisimman täsmällisesti määrittää havaitun vektorin y kuuluvaksi oikeaan luokkaan. Määritellään myös komplementtifunktio i^c siten että $i^c = 1$ jos $i = 2$ ja muulloin $i^c = 2$.

Oletetaan lisäksi, että havaintovektori y noudattaa multinormaalijakaumaa $y \sim N(\mu_i, V_i)$, jossa μ_i on luokkaan i kuuluvien havaintojen keskiarvo ja V_i niiden kovarianssimatriisi. Merkitään π_i:llä prioritodennäköisyyksiä $P\{J = i\}$, $i = 1,2$. Posterioritodennäköisyys sille, että havainto y kuuluu luokkaan 1 tai 2 saadaan Bayesin kaavasta:

$$P\{J = i \mid Y = y\} = \frac{\pi_i \times f(y \mid J = i)}{\pi_1 \times (y \mid J = i) + \pi_2 \times (y \mid J = i^c)}, \qquad (1)$$

jossa $f(y|J=i)$:llä merkitään y:n tiheysfunktiota, annettuna luokka *i*. *Gauss-Bayesilainen luokittelija* määrittää havaintovektorin *y* luokkaan *i*, *g(y)* = *i* jos ja vain jos

$$P\{J=i|Y=y\}>\frac{1}{2}. \tag{2}$$

Edellä määritelty luokittelija on luonteeltaan staattinen, mutta aikasarjojen ollessa kyseessä normaalisuusoletus saa muodon

$$Y_t \sim N(\mu_{J_t}, V_{J_t}) . \tag{3}$$

Olkoon nyt J_t ensimmäisen asteen homogeeninen Markov-ketju, jossa dataa generoivalla prosessilla on kaksi tilaa, joita ei suoraan voida havaita. Kummassakin tilassa havaittujen havaintojen todennäköisyys liittyy jompaankumpaan yllä esitetyistä multinormaalijakauman tiheysfunktioista. Markovin ketju määrittää siirtymisen tilasta toiseen. Markovin todennäköisyyttä sille, että tilaa *i* seuraa tila *j*, merkitään seuraavasti:

$$p_{ij} = P\{J_t = j|J_{t-1} = i\} \tag{4}$$

Mallin määrittävät nyt kumpaankin tilaan liittyvät multinormaalijakaumat (3) ja matriisi P, joka sisältää todennäköisyydet (4). Mallin parametrijoukko on $\Theta = \{\mu_1, \mu_2, V_1, V_2, p_{11}, p_{12}\}$.
Hetkellä t-1 posterioritodennäköisyydet voidaan kirjoittaa:

$$p_i(t-1) = P\{J_{t-1} = i|Y_{t-1} = y_{t-1}\} .$$

Prioritodennäköisyydet tiloille 1 ja 2 hetkellä t voidaan ilmaista Markov-todennäköisyyksien ja edellisten posterioritodennäköisyyksien avulla:

$$\pi_i(t) = p_{1i} \times p_1(t-1) + p_{2i} \times p_2(t-1)$$

ja posterioritodennäköisyydet tietylle havaintovektorille y_t hetkellä *t* ovat

$$p_i(t) = \frac{\pi_i(t) \times f(y_t|J_t = i)}{\sum_{j=1}^{2} \pi_j(t) \times f(y_t|J_t = j)}, \tag{5}$$

jossa $f(y_t|J_t = i)$ on multinormaalijakauman tiheysfunktio y_t:ssä ehdolla

$J_t = i$.

Oletukset (3) ja (4), posterioritodennäköisyydet (5) sekä päätössääntö

$$P\{J_t = i \mid Y_t = y_t\} > \frac{1}{2} \qquad (2')$$

määrittelevät Gauss-Bayesilaisen luokittelijan dynaamisen version, jota Koskinen ja Öller kutsuvat *Markov-Bayesilaiseksi luokittelijaksi.*

Kun Markov-Bayesilaisen luokittelijan parametrijoukko Θ on tunnettu, havaintojen luokitteluun voidaan käyttää esimerkiksi Hamiltonin (1989) esittämää pienimmän neliösumman rekursiivista algoritmia.

KÄÄNNEPISTEINDIKAATTORIN MUODOSTAMINEN

Seuraavaksi tarkoituksena on konstruoida käännepisteitä ennakoiva todennäköisyystyyppinen indikaattori, jota kutsumme käännepisteindikaattoriksi. Mallin valinnassa tarkastelemme indikaattorin suoriutumiskykyä käännepisteiden ennakoinnissa estimointiperiodin ulkopuolella (havaintoaikajakson viimeiset kahdeksan vuotta 1995-2002) ja mallin hyvyyden kriteerinä käytämme vääriin ennakointisignaaleihin pohjautuvan tappiofunktion minimoimista. Koskinen ja Öller perustelevat tämän lähestymistavan käyttöä tavanomaisemman suurimman uskottavuuden menetelmän sijaan, vaikka mallin valinnassa ei tässä voidakaan käyttää esimerkiksi tavanomaisia Schwarzin tai Akaiken informaatiokriteereitä.

Mallin rakentaminen tapahtuu kolmessa vaiheessa, jossa ennakointidatana käytettävien havaintojen suodatinfunktion sekä Markov-todennäköisyyksien parametriestimoinnit tehdään grid-search -etsintämenetelmällä ja iteroimalla, kunnes valitun tappiofunktion minimoivat parametrit löytyvät.[1]

Mallirakentamisen alkuoletuksena on, että aikasarjan (tässä teollisuustuotannon trendisarjan) suhdannekehitys voidaan jakaa kahteen vaiheeseen tai tilaan – kasvuun ja taantumaan – jotka on määritelty yksinkertaisesti em. merkkisäännön mukaan.

Merkitään nyt J_t:llä suhdannevaihetta kuvaavaa aikasarjaa. Tehtävänä on valita malli, jonka tarkoitus on mahdollisimman hyvin ennustaa J_{t+l}, $l>0$ hetkellä t, missä indeksillä l merkitään mallin ennakkoa neljännesvuosina.

Mallissa käytetään hyväksi ennakollisen aikasarjan y_t suodatusta ja siihen sovellettua Markov-Bayesilaista luokittelijaa. Havainnot jaetaan estimointijaksoon (1975-94) ja testijaksoon (1995-2002). Merkitään estimointijak-

soa nyt $\{y_0, y_1, \ldots, y_{T+l-1}\}$:lla. Merkitään maailmantilojen (kasvu tai taantuma) todennäköisyysestimaattien virhettä e_t:llä:

$$e_t = \hat{P}\{J_{t+l} = i | y_t\} - \delta(J_{t+l}, i), \tag{6}$$

jossa Kroneckerin funktio $\delta(J_{t+l}, i)$ saa arvon 1 jos, $J_{t+l} = i$, ja arvon nolla muulloin. Jos nyt kasvujaksoa merkitään indeksillä $i = 1$ ja taantumaa $i = 0$, kuvaa $\hat{P}\{J_{t+l} = i | y_t\}$ siis todennäköisyysarviota sille, että talous on l periodin kuluttua kasvussa.

Mallin rakentamisessa ja estimoinnissa käytettävän tappiofunktion valinta ei ole suinkaan yksiselitteistä. Koska ennakointi-indikaattorin muodostamisen tavoitteena on väärien käännepistesignaalien minimointi, on luonnollista käyttää tappiofunktiona virheellisten käännepistesignaalien lukumäärää tai suhteellista osuutta. Käännepisteet ovat kuitenkin käytännössä varsin harvinaisia, eikä niitä osu pitkäänkään havaintoaineistoon välttämättä kovin monia. Tämän vuoksi apuna voidaan käyttää esimerkiksi todennäköisyysestimaatin keskineliövirhettä. Merkitään nyt

$$L_1 = \frac{1}{T} \#\{y_t | g(y_t) \neq J_{t+l}\} \text{ ja} \tag{7}$$

$$L_2 = \frac{1}{T} \sum_t e_t^2 . \tag{8}$$

L_1 kuvaa nyt mallin antamien virhehavaintojen suhteellista osuutta (notaatio # tarkoittaa lukumäärää) ja L_2 mallin antamien todennäköisyysestimaattien keskineliövirhettä. Minimoitava tappiofunktio muodostetaan L_1 :n ja L_2 :n painotettuna keskiarvona:

$$L = w \times L_1 + (1-w) \times L_2 , \text{ jossa } w \in [0,1] . \tag{9}$$

Painokertoimen w valinta riippuu siitä, kumpaa tappiofunktion osaa halutaan painottaa. Talouspoliittisen päätöksenteon kannalta käännepistevirheiden minimointi (L_1) olisi tärkeää, minkä vuoksi olemme valinneet w:n arvoksi 2/3 (Koskinen ja Öller käyttivät arvoa 1/3).

Mallin estimointi tapahtuu kolmen vaiheen toistolla:

Vaihe 1: *Havaintojen suodatus.* Havainnot $\{y_0, y_1, \ldots, y_T\}$ suodatetaan suodattimella F. Merkitään suodatettuja havaintoja $\tilde{y}_t = F(y_t)$

Vaihe 2: *Havaintojen jako kahteen joukkoon ja parametrien estimointi.* Suhdannevaihtelua kuvaavan vertailuaikasarjan (tässä teollisuustuotannon esti-

moitu trendi) havainnot I = {0,1,...,T} jaetaan kahteen kasvu- ja taantumaperiodiin:

$$I(i) = \{t \in I \,|\, J_{t+l} = i\} \ , \ i = 1,2.$$

Tämän jälkeen estimoidaan keskiarvot ja kovarianssimatriisit kummallekin joukolle:

$$\hat{\mu}_i = \frac{1}{\#\{I(i)\}} \sum_{t \in I(i)} \tilde{y}_t$$

$$\hat{V}_i = \frac{1}{\#\{I(i)\}} \sum_{t \in I(i)} (\tilde{y}_t - \hat{\mu}_i)(\tilde{y}_t - \hat{\mu}_i)' .$$

Vaihe 3: *Markov-todennäköisyyksien estimointi.* Vaiheessa 1 suodatettujen havaintojen sekä vaiheessa 2 estimoitujen keskiarvojen ja kovarianssimatriisien avulla lasketaan estimaatit Markovin todennäköisyyksille Hamiltonin iteratiivisella menetelmällä (Hamilton 1989) ja valitaan se Markovin matriisi P joka minimoi tappiofunktion *L*.

Käyttämämme menetelmä noudattaa muuten Koskisen ja Öllerin vastaavaa, mutta tappiofunktiota minimoitaessa olemme poistaneet sarjan ensimmäisen havainnon eri vaihtoehdoissa sekä tunnuslukujen raportoinneissa. Tämä siksi, että Markovin todennäköisyyksien estimoinnissa käytettävä iteraatioalkuarvo (=0.5) on luonteeltaan *ad hoc*, eikä sitä voida pitää vastaavana todennäköisyysindikaattorina kuin aikasarjan myöhempiä havaintoja.

Alkuarvon 0,5 ottaminen mukaan johtaa yhtäältä tappiofunktion minimoinnissa keskineliövirheosan L_2 perusteettomaan kasvuun. Mallin muusta käyttäytymisestä riippumatta ensimmäisen havainnon neliövirheeksi saataisiin aina $0{,}5^2 = 0{,}25$, ja se vaikuttaisi lopputuloksiin sitä enemmän, mitä lyhyemmästä aikasarjasta on kyse.

Alkuarvon mukaan ottaminen johtaisi toisaalta väärään päätelmään suhdannetilanteesta siinä tapauksessa, että ensimmäisen periodin regiimi on eri kuin laskenta-algoritmin alkuarvosta 0,5 päätössäännön (2′) mukaan tehty johtopäätös. Näin ollen alkuarvo voi vääristää myös virhesignaalien suhteellista osuutta L1:tä.

KÄÄNNEPISTEVIRHEIDEN LUOKITTELUSTA

Käännepistevirheiden suhteellisen osuuden lisäksi niitä voidaan arvioida tarkemminkin. Tämä on relevanttia varsinkin silloin, kun eri tyyppisiin tilanteisiin ja virheisiin liittyvät tappiofunktiot eivät ole symmetrisiä, tai kun en-

nakointi-indikaattorin käyttötarkoituksissa on eroavuuksia käyttäjistä riippuen.

Theil (1958) on jakanut talousennusteissa tehdyt käännepistevirheet kahteen luokkaan: 1) virheisiin, joissa on ennustettu käänne, jota ei kuitenkaan ole tapahtunut sekä 2) virheisiin, joissa toteutunut käännepiste on jätetty ennustamatta. Theilin menetelmän mukaisesti käännepiste-ennusteiden osuvuutta voidaan tarkastella taulukon 12.1 mukaisessa kehikossa, jossa kirjaimet a-d kuvaavat kunkin havaintotyypin lukumäärää aineistossa:

Tyypin 1 käännepistevirheiden lukumäärää kuvaa taulukon luku c ja tyypin 2 käännepistevirheiden lukumäärää luku b. Theil muodostaa lisäksi kaksi kvantitatiivista mittaa, jotka voidaan muodostaa esitetyn taulukon avulla ja joilla voidaan arvioida käännepiste-ennusteiden luotettavuutta.

$$\sigma_1 = \frac{b}{a+b}$$

$$\sigma_2 = \frac{c}{a+c}$$

σ_1 kuvaa ennustamatta jääneiden käänteiden osuutta kaikista käänteistä ja σ_2 väärien käännepiste-ennusteiden osuutta kaikista käännepiste-ennusteista. Luokittelua on käytetty mm. ETLAn eri ennustemuuttujien virheiden käännepisteanalyysissä (Mustonen, 1982).

Theilin käännepistevirheitä kuvaavat tunnusluvut tuovat lisäinformaatiota arvioitaessa Markov-bayesilaisen luokittelijan toimivuutta. Tunnuslukuja on mahdollista kehittää edelleen ottamaan paremmin huomioon käännepisteindikaattoreiden erilaiset käyttötarkoitukset. Kun tarkasteltavan aikasarjan kehitys jaetaan kasvu- ja taantumavaiheisiin, voidaan käännepistevirheet jakaa neljään luokkaan sen mukaan, missä tilassa talous on virheen tapahtuessa. Virheet ovat seuraavantyyppiset:

1. "Väärä taantumasignaali": talous on kasvussa, ja malli ennustaa käännettä taantumaan, joka ei kuitenkaan toteudu.
2. "Puuttuva taantumasignaali": talous on kasvussa ja kääntyy taantumaan, mutta malli ei tätä ennusta.
3. "Väärä kasvusignaali": talous on taantumassa, ja malli ennustaa käännettä kasvuun, joka ei kuitenkaan toteudu.
4. "Puuttuva kasvusignaali": talous on taantumassa ja kääntyy kasvuun, mutta malli ei tätä ennusta.

Aikaisempia merkintöjä käyttäen väärä signaali saadaan, jos $e_t > 0.5$. Erityyppisiä virheitä voi kuvata taulukon 12.2 nelikentällä, jossa miinus- ja plusmerkit kuvaavat kahden peräkkäisen ennustevirheen etumerkkejä. Käännepisteitä ennakoivan mallin toimivuutta tulisi voida arvioida myös tarkastelemalla näiden eri tyyppisten ennustevirheiden osuuksia.[2]

TULOKSET

Seuraavassa raportoimme estimoidut mallit, joissa on käytetty aineistona kausitasoitetun teollisuustuotannon neljännesvuosittaista muutosta sekä TT:n Suhdannebarometrin aikasarjoja. Ensimmäisessä mallissa käytetään suhdannenäkymiä ja tuotanto-odotusta, toisessa suhdannenäkymiä ja tilauskantaa ja kolmannessa suhdannenäkymiä ja vientiodotusta. Vertailun vuoksi on laskettu myös malli nolla, jossa on pelkästään kaksi muuttujaa – teollisuustuotanto ja suhdannenäkymät sekä benchmark-malli, jossa ei käytetä lainkaan suhdannekyselydataa. Benchmark-mallin ainoana selittäjänä on itse teollisuustuotantoaikasarja.

Edellä selostetun estimointirutiinin vaiheessa 1 tehtävä suodatus on käytännössä tehty eksponentiaalisella liukuvalla keskiarvolla, jolloin suodatetut havaintosarjat ovat muotoa $\hat{y}_t = \lambda y_t + (1-\lambda)\hat{y}_{t-1}$. Malleissa ennakoksi on

Taulukko 12.1. Käännepistevirheiden luokittelukehikko

		Ennustettu	
		Käännepiste	Ei käännepistettä
Toteutunut	Käännepiste	a	b
	Ei käännepistettä	c	d

Lähteenä on Theil (1958).

Taulukko 12.2. Käännepisteiden ennustevirheiden nelikenttä

		Talouden tila hetkellä t	
		Kasvu	Taantuma
Talouden tila hetkellä t - 1	Kasvu	- - Väärä taantumasignaali	-+ Puuttuva taantumasignaali
	Taantuma	+- Puuttuva kasvusignaali	++ Väärä kasvusignaali

asetettu 1. vuosineljännes ja vertailun vuoksi on esitetty tulokset samoista malleista niin, että ennakko on nolla.

Tässä olemme etsineet mallia, joka ennustaisi teollisuustuotannon neljännesvuosittaiset käännepisteet yhden neljänneksen ennakolla. Koskisen ja Öllerin sovelluksissa on haettu mallia, jonka vertailuaikasarjana on käytetty teollisuustuotannon vuosimuutokseen perustuvaa vertailuaikasarjaa. Tämän sarjan käänteet tapahtuvat selvästi myöhemmin kuin käyttämämme aikasarjan.

Vertailusarjan kasvu- ja taantumajaksojen sekä käännepisteiden määrittely on tehty teollisuustuotannon trendisarjan perusteella, koska oletamme teollisuustuotannon oleellisten käänteiden löytyvän paremmin sen trendisarjasta *ex post*. Ajantasaseurannassa trendiarvioiden muodostaminen käytetyn X12-mallin avulla antaa kuitenkin helposti vääriä signaaleja juuri käännepisteiden kohdalla. Siten todellisessa ennakointitilanteessa kausitasoitettujen neljännesvuosilukujen käyttö on perustellumpaa, jonka vuoksi käytämme sitä input-sarjana.

Mallin valinnassa olemme vertailleet valittuja malleja niiden tappiofunktioiden perusteella testiperiodilla 1995-2002, johon kuuluu kaksi taantumajaksoa. Lisäksi periodiin sisältyy vuoden 1998 ns. Aasian kriisin jakso, jolloin aggregaattisarjassa ei havaita käänteitä, mutta elektroniikka- ja sähköteollisuudesta puhdistetussa teollisuustuotantosarjassa nähdään lyhyt taantuma (ks. kuvio 12.1).

Yhden vuosineljänneksen ennakolla parhaaksi malliksi osoittautuu M31, jossa ennakoivina muuttujina ovat teollisuustuotannon lisäksi TT:n barometriaineiston suhdannenäkymien ja vientiodotusten saldomuuttujat. Ilman suhdanneindikaattoreita muodostettuun benchmark-malliin liittyvät tappiofunktion arvot ovat sekä estimointi- että testijaksolla selvästi suuremmat kuin malleissa, joissa on mukana yksi tai kaksi suhdanneindikaattoria. Suhdanneindikaattoreiden käyttö käännepiste-ennakoinnissa parantaa siis selvästi ennakointikykyä. Mallissa M31 paras tulos saadaan myös sarjan tasoituskertoimella 1, eli mukaan tulee vain ajankohtaisin havainto.

Testijaksolla parhaiten toiminut malli M31 myöhästyi vuosien 1995-96 ns. minitaantuman ennakoinnissa, sen sijaan vuoteen 2001 ajoittuneen taantuman M31 se ennakoi hyvin. Vuoden 1998 Aasian kriisiin, jolloin laajasti ottaen muut toimialat kuin elektroniikka- ja sähköteollisuus kokivat lyhyen taantuman, se ei myöskään reagoinut; muissa kokeilluissa malleissa tähän jaksoon liittyi varoituksen merkkejä, vaikka nekään eivät ennakoineet taantumaan joutumista.

Mielenkiintoista on myös käännepisteiden ennakointikyky estimointijaksolla. Jos tulkitaan todennäköisyysestimaatin siirtyminen lähemmäksi hälytysarvoa 0.5 heikoksi signaaliksi mahdollisesta muutoksesta, niin 1990-luvun alun suuren laman alkua nyt käsitelleet mallit eivät osanneet ajoittaa, vaikka jonkinlaiset varoitukset olivat nähtävillä: M31:n mukainen todennäköisyysestimaatti laski arvoon 0.7 yksi vuosineljännes ennen käännepistettä.

Myös vuonna 1986 nyt konstruoidut indikaattorit reagoivat voimakkaasti. Suurin osa nyt estimoiduista indikaattoreista oli lähellä taantumahälytystä. Esimerkiksi benchmark-mallina käytetty malli, jossa mukana oli vain teollisuustuotannon havainnot, hälytti selvästi taantumasta. Tuohon ajankohtaan liittyikin tuotannon kasvuvauhdin nopea hidastuminen. Varsinaista laskua tapahtui vain kuitenkin vain yhden vuosineljänneksien ajan, joten kyseessä oli tässä käytetyn taantuma-käsitteen sijaan ns. kasvutaantuma.

Verrattaessa eri pituisten ennakointihorisonttien malleja, päästään nollaennakolla parempiin ennakointimalleihin sekä estimointi- että ennakointiperiodeilla, mikä on odotettavaakin. Testiperiodin perusteella parhaiksi malleiksi osoittautuivat mallit, joissa muuttujina ovat teollisuustuotannon ja suhdannenäkymien lisäksi tuotanto-odotukset (M10) tai tilauskanta (M20). Malli M20 toimi testijaksolla hieman paremmin kuin M10, joka puolestaan oli estimointijaksolla selvästi parempi. On sinällään luontevaa, että eri mallit toimivat eri tavoin eri ennakointihorisonteilla; vientiodotuksen sisältävä malli oli kokeilluista malleista heikoin testijaksolla, joskin estimointijaksolla se antoi muihin verrattuna hyviä tuloksia.

Taulukko 12.3. Käännepisteindikaattorimallien ennakointikyky[3]

	λ_1	λ_2	λ_3	p11	p22	L1(e)	L2(e)	L(e)	L1(t)	L2(t)	L(t)
BM1	0.5	-	-	0.61	0.23	0.113	0.108	0.110	0.158	0.266	0.231
M01	1.0	1.0	-	0.95	0.53	0.068	0.068	0.068	0.118	0.133	0.128
M11	1.0	1.0	0.7	0.88	0.36	0.064	0.054	0.057	0.113	0.167	0.145
M21	0.9	1.0	0.8	0.76	0.07	0.066	0.054	0.058	0.075	0.100	0.092
M31	1.0	1.0	1.0	0.95	0.51	0.043	0.027	**0.032**	0.061	0.067	**0.065**
BM0	0.6	-	-	0.80	0.60	0.063	0.067	0.065	0.076	0.097	0.090
M00	0.6	1.0	-	0.89	0.67	0.050	0.027	0.034	0.051	0.065	0.060
M10	0.9	1.0	0.7	0.96	0.76	0.020	0.013	**0.016**	0.041	0.065	0.057
M20	0.5	0.6	0.9	0.74	0.15	0.034	0.065	0.054	0.034	0.065	**0.054**
M30	0.7	1.0	1.0	0.96	0.83	0.027	0.013	0.018	0.067	0.097	0.087

Kuviossa 12.2 on esimerkki yhden neljänneksen ennakolla parhaiten toimineesta mallista M31. Kuvion yläosa liittyy estimointijaksoon ja alaosa testijaksoon, jossa on käytetty estimointijakson parametriarvoja. Yhtenäinen viiva kuviossa kuvaa viitesarjan kasvujaksoja (arvo = 1) sekä taantumajaksoja (arvo = 0), ja todennäköisyysestimaatit on esitetty katkoviivalla. Hälytys käänteestä tapahtuu, kun todennäköisyysestimaatti siirtyy viitearvon 0.5 toiselle puolelle joko alhaalta (kasvukäänne) tai ylhäältä (taantumakäänne).

YHTEENVETO JA JOHTOPÄÄTÖKSET

Talouden käännepisteiden ajantasainen ennakointi on yritysten suunnittelun, nopean reagoinnin ja muutosten ajoituksen kannalta usein tärkeämpää kuin vuositason muutosten suhdanneseuranta. Myös talouspolitiikan tekemisen kannalta taantuma- ja kasvujaksojen oikealla ennakoinnilla on merkityksensä. Lyhyen ajan – kuukausi- ja vuosineljännespohjaisen – seurannassa kyselyaineistojen hyväksikäyttö on yleistynyt viime vuosina kaikkialla. Suomessa Teollisuuden ja Työnantajain Keskusliiton keräämä suhdannebarometriaineisto on saavuttanut tässä suhteessa vankan aseman. Tässä artikkelissa olemme käsitelleet suhdannekyselyaineiston hyväksikäyttöä teollisuustuotannon suhdannekäänteiden ennakoinnissa.

Olemme pyrkineet konstruoimaan teollisuustuotannon käännepisteistä ennakoivan käännepisteindikaattorin, joka perustuu ns. Markov-Bayesilaiseen luokittelijaan ja johon yhdistetään TT:n suhdannekyselyaineiston ennakoivia muuttujia. Mallitus perustuu Koskisen ja Öllerin (1998) vastaavaan indikaattoriin, jota he ovat soveltaneet Ruotsin aineistoon.

Suhdannekyselyaineiston relevanssia perustelemme sen nopealla saatavuudella, pitkään jatkuneella perinteellä aineiston keräämisessä ja hyödyntämisessä sekä aineiston ennakoivalla luonteella. Esittämiemme tulosten perusteella vaikuttaa ilmeiseltä, että suhdannekyselyaineiston käyttö parantaa ennakointikykyä verrattaessa sellaiseen malliin, jossa tätä informaatiota ei hyödynnetä. Koskinen ja Öller ovat todenneet omilla aineistoillaan käytetyn menetelmän hyödyllisyyden myös moniin yksinkertaisiin ja ns. naiiveihin malleihin verratessa.

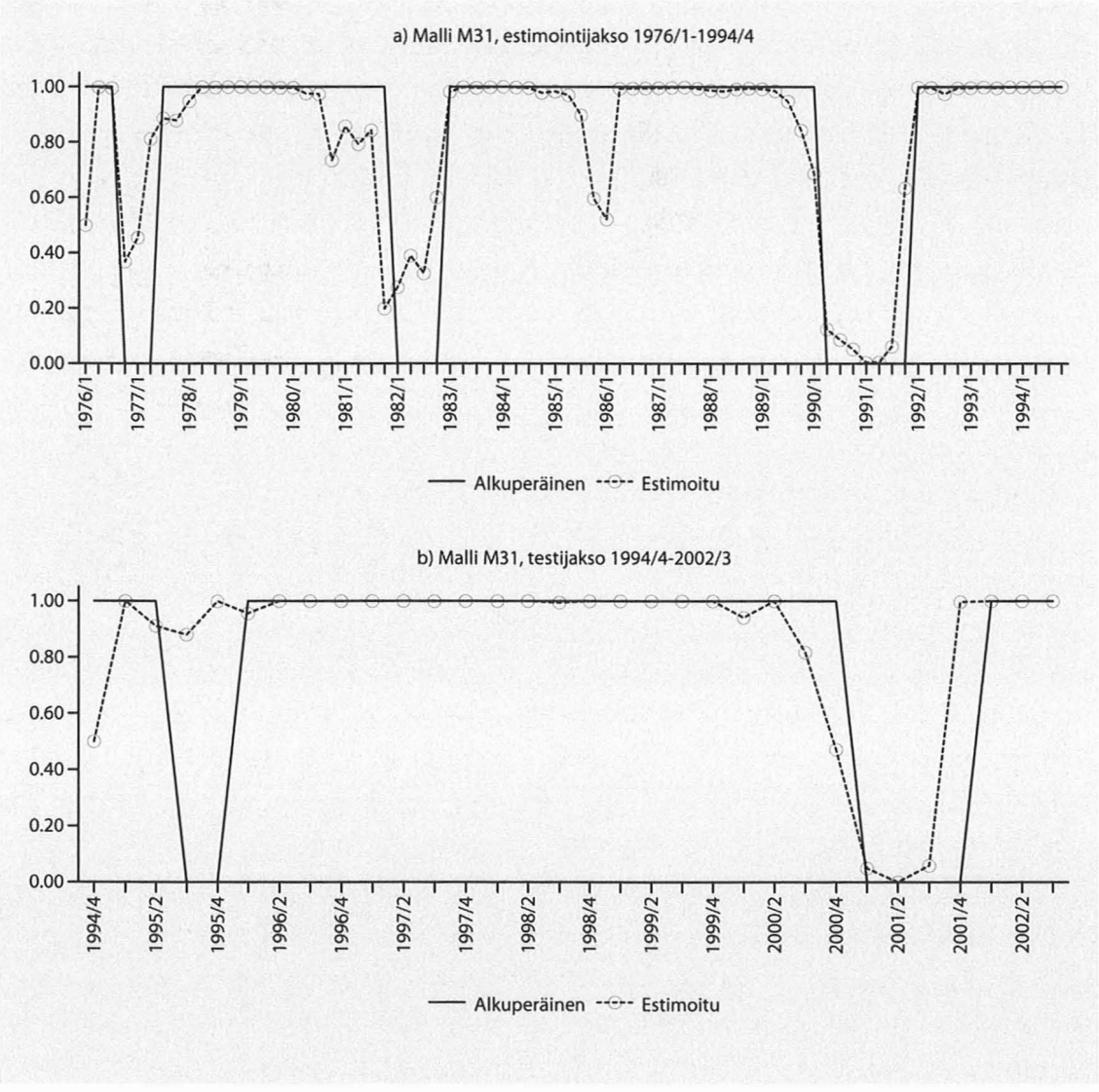
Käytetty datalähtöinen mallitustapa sallii erityyppisten ennakointitavoitteiden valinnan ja tappiofunktioiden spesifioinnin joustavasti käyttötarpeiden mukaan. Käyttäjästä ja käyttötarkoituksesta riippuen on nimittäin mahdollista, että halutaan keskittyä esimerkiksi taantumien/kasvujaksojen ennakointiin silläkin riskillä, että väärät hälytykset ovat yleisiä. Tällöin voi-

daan käyttää tappiofunktiota, joka rankaisee tekemättä jääneistä taantumahälytyksistä eri tavalla kuin turhista hälytyksistä.

Parhaaksi yhden neljänneksen ennakon käännepisteindikaattoriksi osoittautui testiperiodilla 1995-2002 malli, jossa muuttujina käytettiin teollisuustuotannon kausitasoitettujen neljännesvuosimuutosten lisäksi barometrimuuttujista suhdannenäkymien ja vientiodotusten saldoa. Tämä on aikaisempien selvitysten ja odotusten kanssa yhteensopivaa teollisuutemme rakenteen ja suuren vientiriippuvuuden vuoksi.

Nyt esitetyt mallit ja indikaattorit ovat vielä monelta osin alustavia, mutta tuovat mielestämme mielenkiintoista ja tärkeää lisäinformaatiota talouden seurantaan ja ennakointiin. Jatkon kannalta käytetty menetelmä ja sen

Kuvio 12.2. Teollisuustuotannon regiimit ja käännepisteindikaattori



Teollisuustuotannon regiimissä 1 viittaa kasvuun ja 0 taantumaan.

soveltamiseen kehitetty estimointi- ja laskentakehikko tarjoavat runsaasti mahdollisuuksia käännepisteindikaattoreiden tarkastelulle myös laajemman datan pohjalta. Myös toimialakohtaisten mallien rakentaminen on yksi potentiaalinen sovelluskohde, koska barometriaineiston toimialoittaisessa tarkastelussa on havaittu selviä eroja toimialakohtaisten vastausten ennakointikyvyssä (Urrila 1999).

KIRJALLISUUTTA

Bergström, R. (1993): The full trichotomous scale compared with net balances in qualitative business survey data - experiences from the Swedish business tendency surveys. Konjunkturinstitutet Working Papers No. 30a. Konjunkturinstitutet, Stockholm.

Hamilton, J.D. (1989): A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica, 57, 357-384.

Kangasniemi, J. ja Tsupari, P. (1997): Manufacturing Production and the Business Tendency Survey in Finland: The Concept of Business Outlook. Teoksessa Oppenländer, K.H. - Poser, G. (toim.): Social and Structural Change - Consequences for Business Cycle Surveys. Selected Papers presented at the 23rd CIRET Conference, Helsinki, 1997. Ashgate Publishing, Aldershot.

Koskinen, L. ja Öller, L.-E. (1998): A Classifying Procedure for Signalling Turning Points. Research Report, National Institute of Economic Research. Konjukturinstitutet, Ekonomiska rådet, Stockholm, Sweden.

Loungani, P. (2000): How Accurate Are Private Sector Forecasts? Cross-Country Evidence from Consensus Forecasts of Output Growth. IMF working paper WP/00/77.

Mustonen J. (1982): ETLAn vuosille 1971 – 1980 laatimien suhdanne-ennusteiden virheanalyysi, ETLA C 25, Helsinki.

Neftçi, S. N. (1982): Optimal Prediction of Cyclical Downturns. Journal of Economic Dynamics and Control 4 (1982).

Takala, K. ja Tsupari, P. (1999): The Predictive Power of Finnish Business Survey Among Industrial Sectors. Teoksessa Oppenländer, K. H. - Poser, G. - Schips, B. (toim.): Use of Survey Data for Industry, Research and Economic Policy. Selected papers presented at the 24th CIRET Conference, Wellington, New Zealand, 1999. Ashgate Publishing, Aldershot.

Teräsvirta, T. (1985): Teollisuustuotannon volyymin ennustaminen Suhdannebarometrin avulla. ETLA B:43. Taloustieto Oy. Helsinki.

Theil, H. (1958): Economic Forecast and Policy. Amsterdam.

Urrila, P. (1999): Teollisuuden ja Työnantajain Keskusliiton Suhdannebarometri sekä teollisuustuotannon muutos eri toimialoilla. Teollisuuden ja Työnantajain Keskusliitto, Helsinki.

Urrila, P. (2001): Suhdanneindikaattoreiden käyttö talouskehityksen seurannassa. ETLAn keskustelualoitteita 780. Elinkeinoelämän Tutkimuslaitos, Helsinki.

Öller, L. E. (1992): Eliciting Turning Point Warnings from Business Surveys. Konjunkturinstitutet Working Papers No. 17. Konjunkturinstitutet, Stockholm.

LOPPUVIITTEET

[1] Kiitämme Lasse Koskista estimoinnissa käytettävien ohjelmanosien ja algoritmin saamisesta käyttöömme sekä Jouko Kangasniemeä menetelmän APL-pohjaisen sovelluksen tekemisestä sekä laskentamenetelmän ja raportoinnin kehittämisestä tarpeisiimme.

[2] TT:ssä käytetty käännepisteindikaattoreiden estimointiohjelma sallii minimoitavan tappiofunktion konstruoimisen yllä olevan taulukon mukaisten merkkisääntöjen avulla siten, että erityyppisten signaalien painoarvoa voidaan painottaa eri tavoin, käyttötarkoituksesta riippuen.

[3] Taulukossa ennakointineljänneksiä (0 tai 1) kuvaa mallilyhenteen kolmas kirjain 0 tai 1. Malli BM on "benchmark-malli", jossa on mukana vain teollisuustuotanto. Mallissa M0 on teollisuustuotannon lisäksi suhdannenäkymät, ja malleissa M1, M2 ja M3 on edellisten lisäksi vastaavasti seuraavat barometrimuuttujat: tuotanto-odotus, tilauskanta sekä vientiodotus. λ_1 on teollisuustuotantosarjan tasoituskerroin, λ_2 suhdannenäkymät-sarjan tasoituskerroin ja λ_3 kunkin mallin kolmannen muuttujan tasoituskerroin. Markovin todennäköisyydet ovat sarakkeissa *p11* ja *p22*. L1, L2 ja L ovat kaavan (9) mukaisten tappiofunktioiden arvot, *(e)* viittaa estimointijaksoon 1975-94 ja *(t)* viittaa testijaksoon 1994-2002. Käytetyssä grid-search -estimoinnissa tasoitusparametrien etsintävälinä oli 0.1 ja Markovin todennäköisyyksien etsintäväli oli 0.01.

13. Analysing Short Heterogeneous Panel Data: The Demand for Labour in Finnish Firms

Markku Rahiala* and Tapani Kovalainen**

Abstract:
In this article we study the analysis of short firm-level panel data, when the panel is heterogeneous in the sense that the behaviour of different firms cannot be mimicked by one and the same model with constant coefficients. In order to avoid identification problems, the parameters will then have to be interpreted as random variables attaining different values for each firm. The encountered estimation problems are methodologically especially challenging, if ARX- type models including lagged endogenous variables are involved. However, if the purpose of the study is to estimate some non-linear functions of the coefficients, like for instance long-term elasticities, it will be utmost important to take account of the heterogeneity and to estimate the elasticities individually.

The use of random coefficient ARX- models is demonstrated by modelling the firm-specific demand for labour in Finnish companies. The data show very marked heterogeneity indeed.

INTRODUCTION

Dynamic ARX- models for panel data have been studied fairly extensively in the econometric literature during the last 15 years. These models have been used for instance to study production functions with autocorrelated errors (Blundell and Bond, 2000) and to study labour demand both on the firm level (cf. e.g. Arellano and Bond, 1991 or Blundell and Bond, 1998) and on the branch level (cf. e.g. Pesaran and Smith, 1995). Most of the available literature is

* Markku Rahiala is Professor of Econometrics at the University of Oulu
** Tapani Kovalainen is Lecturer of Economics at the University of Lapland

concentrated on models with fixed coefficients and their estimation by the so-called GMM- method (cf. e.g. Arellano and Bond, 1998). On the other hand, Pesaran and Smith (1995) and Pesaran and Zhao (1999) have considered heterogeneous panel data, where the coefficients of the ARX- models vary across branches. Hoogstrate *et al.* (2000) have also looked at a technically similar problem, but all of them assume the time series to be long enough to allow for separate estimation of the individual coefficient vectors. However, panel data sets on the firm level are usually too short to allow for estimation of the whole model for each firm, whereas longer time series might be available on more aggregated levels.

If the panel turns out to be heterogeneous, i.e. if different firms seem to behave differently, it would of course be unrealistic to mimic their behaviour with one and the same model. According to our experience, quite many panel data sets actually show heterogeneity of this kind. If the purpose of the study is to estimate some long-term elasticities or some other non-linear derivatives of the original parameters, it will be important to take the heterogeneity into account and to estimate the elasticities firm by firm. The expectations of the long-term effects can then be estimated by taking averages over the corresponding firm-specific estimates.

There are lots of statistical difficulties, however, in modelling heterogeneous panel data. If we would take the firm-specific coefficients as ordinary parameters, we would face the well-known incidental parameter problem (cf. e.g. Lehman, 1983, p. 450) leading to inconsistent ML- estimators. Furthermore, the firm-specific parameters would not even be identifiable in case of short time series. This is why the only available option is to interpret the coefficients as random variables attaining different values for different firms. In order to solve the identification problem due to the shortness of the data set, we will have to postulate the shape of the joint distribution of these random coefficients.

We will discuss the statistical problems in Section 2. The demand for labour in Finnish firms will be studied in Section 3.

DYNAMIC MODELS WITH RANDOM COEFFICIENTS

ESTIMATION AND TESTING

Let y_{it} denote the t:th measurement of the dependent variable from firm i , $t = 1,...,T_i$, $i = 1,...,N$. Further, let $X_{it} = (x_{i1t} \quad ... \quad x_{imt})'$ be an m- dimensional vector of exogenous explanatory variables. We will study the following model framework for heterogeneous panel data:

$$\begin{aligned} y_{it} &= \mu_i + \phi_{i1} y_{i,t-1} + ... + \phi_{ip} y_{i,t-p} + \beta_i' X_{it} + \varepsilon_{it} \quad , \\ \varepsilon_{it} &\sim \text{NID}(0, \sigma^2) \qquad t = p+1, ..., T_i \quad , \end{aligned} \tag{1}$$

where the parameters $\theta_i = (\mu_i \quad \phi_{i1} \quad ... \quad \phi_{ip} \quad \beta_i')'$ are allowed to vary over firms. The processes $\{\varepsilon_{it}\}$ and $\{X_{it}\}$ are assumed independent of each other, which rules out any instantaneous feedback relations from y_{it} to X_{it} . On the other hand, the vector of explanatory variables X_{it} may include various group dummies or time dummies that do not vary across firms. Neither $\{y_{it}\}$ nor $\{X_{it}\}$ are assumed stationary. This is important, because stationarity can usually be achieved only through differencing the observations one or several times. After differencing, however, it might be unrealistic to assume $\{\varepsilon_{it}\}$ to be white noise. In the sequel, we will further denote $X_{it}^* = (1 \quad y_{i,t-1} \quad .... \quad y_{i,t-p} \quad X_{it}')'$ and $S_i = (y_{i,1} \quad .... \quad y_{i,p})'$.

Quite often only some of the components of θ_i will be allowed to vary over firms and will be assumed stochastic. This is why we write θ_i into the form

$$\theta_i = Q\zeta_i^* + \theta^* \tag{2}$$

where Q is a $(p + m + 1)$ x r- matrix of zeros and ones with just one non-zero element in each column. The vector of individual random effects ζ_i^* will be assumed independent of the innovation shocks ε_{it} . Because the observed processes $\{y_{it}\}$ have been going on for some time before the start of the observation period, it will be necessary to allow the initial values S_i to depend on the firm-specific random effects ζ_i^* . This is why we assume ζ_i^* to have a joint multinormal distribution with the initial values S_i ,

$$\begin{pmatrix} \zeta_i^* \\ S_i \end{pmatrix} \sim \text{NID}_{r+p}\left(\begin{pmatrix} 0 \\ \mu^* \end{pmatrix} , \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}' & \Omega_{22} \end{pmatrix} \right) \quad . \tag{3}$$

Assumption (3) immediately implies that

$$\zeta_i^* \mid S_i \sim \text{N}_r\left(B(S_i - \mu^*) \ , \ \Omega_o\right)$$

where $B = \Omega_{12}\Omega_{22}^{-1}$ and $\Omega_o = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{12}'$. Furthermore, the corresponding conditional distribution of $\zeta_i = \zeta_i^* - B(S_i - \mu^*)$ will be

$$\zeta_i \mid S_i \sim \mathrm{N}_r\,(0\ ,\ \Omega_o)$$

which means that ζ_i will be independent of S_i. According to our notation, the model $(1)+(2)+(3)$ can be written as $Y_{it} = X_{it}^{*\prime}\,[Q(\zeta_i + B(S_i - \mu^*)) + \theta^*] + \varepsilon_{it}$. We can now write the whole model conditionally on S_i into the form

$$\begin{aligned}
y_{it} &= \tilde{X}_{it}'\theta + Z_{it}'\zeta_i + \varepsilon_{it} \quad , \\
\varepsilon_{it} &\sim \mathrm{NID}(0,\sigma^2) \quad , \quad \{\varepsilon_{it}\} \perp\!\!\!\perp \{\zeta_i\} \quad , \\
\zeta_i &\sim \mathrm{NID}_r(0,\Omega_o) \quad , \quad \zeta_i \perp\!\!\!\perp S_i \\
t &= p+1,...,T_i \quad , \quad i = 1,...,N \quad ,
\end{aligned} \tag{4}$$

where

$$Z_{it} = Q'X_{it}^* \quad , \quad \tilde{X}_{it} = \begin{pmatrix} X_{it}^* \\ S_i \otimes Z_{it} \end{pmatrix} \quad , \quad \theta = \begin{pmatrix} \theta^* - QB\mu^* \\ \mathrm{vec}\,(B) \end{pmatrix} \quad ,$$

$S_i \otimes Z_{it}$ denotes the Kronecker product $S_i \otimes Z_{it} = (\,y_{i1}Z_{it}' \quad .... \quad y_{ip}Z_{it}'\,)'$ and $\mathrm{vec}\,(B)$ denotes the vectorised version of the r x p- matrix B.

By stacking all observations from the same individual below each other, model (4) can be formulated as

$$\begin{aligned}
Y_i &= X_i\theta + Z_i\zeta_i + \varepsilon_i \quad , \\
\varepsilon_i &\sim \mathrm{NID}_{T_i-p}(0,\sigma^2 I) \quad , \quad \{\varepsilon_i\} \perp\!\!\!\perp \{\zeta_i\} \quad , \\
\zeta_i &\sim \mathrm{NID}_r(0,\Omega_o) \quad , \quad \zeta_i \perp\!\!\!\perp S_i \quad , \quad i = 1,...,N \quad ,
\end{aligned} \tag{5}$$

where

$$Y_i = \begin{pmatrix} y_{i,p+1} \\ \cdot \\ \cdot \\ y_{i,T_i} \end{pmatrix} \quad , \quad Z_i = \begin{pmatrix} Z_{i,p+1}' \\ \cdot \\ \cdot \\ Z_{i,T_i}' \end{pmatrix} \quad , \quad X_i = \begin{pmatrix} \tilde{X}_{i,p+1}' \\ \cdot \\ \cdot \\ \tilde{X}_{i,T_i}' \end{pmatrix} \quad , \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i,p+1} \\ \cdot \\ \cdot \\ \varepsilon_{i,T_i} \end{pmatrix} \quad .$$

The likelihood- function defined by Y_i conditionally on S_i and ζ_i will be of the form

$$L_{Y_i|S_i,\zeta_i}(\theta,\Omega_o,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{T_i-p} e^{-\frac{1}{2\sigma^2}(Y_i - X_i\theta - Z_i\zeta_i)'(Y_i - X_i\theta - Z_i\zeta_i)} \tag{6}$$

By taking expectations with respect to ζ_i we get the actual likelihood conditional only on S_i

$$L_{Y_i|S_i}(\theta, \Omega_o, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{T_i - p} e^{-\frac{1}{2\sigma^2}(Y_i - X_i\theta)'(Y_i - X_i\theta)} E_{\zeta_i}\left(e^{\zeta_i' A_i \zeta_i + B_i'\zeta_i}\right) \tag{7}$$

where $A_i = -\frac{1}{2\sigma^2} Z_i' Z_i$ and $B_i = \frac{1}{\sigma^2} Z_i'(Y_i - X_i\theta)$. By straightforward calculations and by repeated use of the Sherman-Morrison-Woodbury- formula (cf. e.g. Henderson and Searle, 1981), the likelihood (7) can be written in the form

$$L_{Y_i|S_i}(\theta, \Omega_o, \sigma^2) \propto \frac{1}{\sqrt{\det(V_i)}} \cdot e^{-\frac{1}{2}(Y_i - X_i\theta)' V_i^{-1} (Y_i - X_i\theta)} \tag{8}$$

where

$$V_i = \sigma^2 (I + Z_i \Omega_o Z_i') \quad .$$

Computationally, this is exactly the standard form of a likelihood corresponding to a mixed linear model (cf. e.g. Diggle *et al.*, 1994, pp. 80-81), but the y_{it}- observations are incorporated in the X_i- and Z_i- matrices as well. As shown by Rahiala (1999), this implies that the information matrix will not be block diagonal and that the covariance matrix of the ML- estimators $\widehat{\theta}$ can not be estimated as in the case of regular mixed models.

Assume now that the number of individuals N with $T_i > p + r$ increases. Rahiala (1999) has shown that the likelihood (8) satisfies the necessary regularity conditions (cf. e.g. Lehman, 1983, p. 429 or Self and Liang, 1987) guaranteeing the asymptotic normality of the ML- estimators with increasing N, if the true value of the parameter is an inner point of the parameter space, if each Z_i matrix is of full rank with probability 1 and if the moment matrix of the exogenous variables $(1/N)\sum X_i' X_i$ tends to a finite positive definite limit with increasing N.

All individuals with $T_i > p$ contain some information on the parameters and are worth taking into account in the estimation, but in order to make use of the asymptotic normality of the ML- estimators, we have to assume that the number of individuals with $T_i > p + r$ is large.

The covariance matrix of the asymptotic normal distribution of the ML-estimator $\widehat{\theta}$ can be evaluated by inverting the outer product estimator of the whole information matrix as shown by Rahiala (1999). He also gave explicit formulae for the calculation of the score vectors needed in the outer product estimator.

Restrictions on the elements of Ω_o that put some of the parameters on the boundary of the parameter space are somewhat tricky to test. They can in fact be tested by the usual LR- statistics, but the asymptotic ditributions of the test statistics will be complicated, non-standard mixtures of several χ^2-distributions. See Self and Liang (1987), Stram and Lee (1994, 1995) or Rahiala (1999) for more details.

To make the multinormality assumption (3) of the random effects more acceptable, it is often advisable to get rid of the intercept terms μ_i by differencing the model into the form

$$\begin{aligned} \Delta y_{it} &= \phi_{i1}\Delta y_{i,t-1} + ... + \phi_{ip}\Delta y_{i,t-p} + \beta_i'\Delta X_{it} + \Delta\varepsilon_{it} \quad , \\ \varepsilon_{it} &\sim NID(0,\sigma^2) \quad , \quad t = p+2, ..., T_i \quad , \end{aligned} \tag{9}$$

where Δ stands for the difference operator and

$$\begin{aligned} &\theta_i = (\,\phi_{i1} \quad ... \quad \phi_{ip} \quad \beta_i'\,)' = Q(\zeta_i + B(S_i - \mu^*)) + \theta^* \quad , \\ &\zeta_i \sim NID_r(0,\Omega_o) \quad , \quad \zeta_i \perp\!\!\!\perp S_i \quad , \quad \theta^* = E\,(\theta_i) \quad , \quad i = 1, ..., N \end{aligned}$$

just like before. We now define the $(T_i - p - 1)$ x $(T_i - p)$- matrix D_i as

$$D_i = \begin{pmatrix} 1 & -1 & 0 & .... & 0 & 0 \\ 0 & 1 & -1 & .... & 0 & 0 \\ . & . & . & .... & . & . \\ . & . & . & .... & . & . \\ 0 & 0 & . & .... & 1 & -1 \end{pmatrix} \quad .$$

By stacking the consequtive observations from each firm below each other, we can then write model (9) into a more compact form

$$\begin{aligned} D_iY_i \;&= (\,\Delta y_{i,p+2} \quad .... \quad \Delta y_{i,T_i}\,)' = D_iX_i^o\theta + D_iZ_i\zeta_i + D_i\varepsilon_i \quad , \\ D_i\varepsilon_i \;&\sim \mathrm{NID}_{T_i-p-1}(\,0\;,\;\sigma^2H_i\,) \quad , \quad \{D_i\varepsilon_i\} \perp\!\!\!\perp \{\zeta_i\} \quad , \\ \zeta_i \;&\sim \mathrm{NID}_r(0,\Omega_o) \quad , \quad \zeta_i \perp\!\!\!\perp S_i \quad , \quad i = 1, ..., N \quad , \end{aligned} \tag{10}$$

where $H_i = D_iD_i'$ and $X_i^o = X_i\begin{pmatrix} 0 \\ I \end{pmatrix}$ is just like X_i, but with the first column of ones deleted.

It will then be obvious that the likelihood function, defined by D_iY_i conditionally on S_i will be of the form

$$L_{D_iY_i|S_i}(\theta,\Omega_o,\sigma^2) \propto \frac{1}{\sqrt{\det(V_i)}} \cdot e^{-\frac{1}{2}(D_iY_i - D_iX_i^o\theta)'V_i^{-1}(D_iY_i - D_iX_i^o\theta)} \tag{11}$$

where

$$V_i = \sigma^2(H_i + D_iZ_i\Omega_oZ_i'D_i') \quad .$$

Estimation of the remaining parameters and tests of hypotheses can be based on this likelihood exactly as explained in section 2.1.

APPLICATION TO FINNISH EMPLOYMENT EQUATIONS

As an example of a clearly heterogeneous set of panel data we consider the employment in 260 Finnish companies during the years 1986-1998. The number of available observations per company varies between 7 and 13. The companies represent 16 different branches. The original data set contained information on the balancing of accounts of 500 largest firms in Finland, but we have discarded all firms with less than 100 employees or with less than 7 comparable observations. The original data were collected by Talouselämä magazine and have been later maintained by ETLA. The branch-level data have been extracted from the National Accounts, Statistics Finland.

Our basic model is based on standard theory of labour demand in case of imperfect competition in the firm's product market as presented for instance in Layard *et al.* (1991, Ch. 7), in Hamermesh (1993, Ch. 2) and in Nickell (1986). We consider a firm maximising its profits

$$\Pi_t = P(Q_t, A_t) \cdot F(L_t, K_t) - W_t L_t$$

with respect to labour L_t. Here W_t denotes the average wages, K_t the gross capital in the firm, Q_t denotes the volume of the aggregate output of the whole branch and A_t denotes an exogenous demand factor. The production function has been denoted by F and the prices of the products by P. We assume the conjecture function measuring the competitors' expected reactions to changes in the output of the firm to be of the form

$$\frac{d(Q_t^e - q_t)}{dq_t} = c - 1 = \text{constant} \quad , \tag{12}$$

where Q_t^e stands for the expected value of Q_t and q_t stands for the output of the firm.

The first order optimality conditions lead to a demand function for labour

$$L_t^* = L\left(m_t \cdot \frac{W_t}{P_t}, K_t\right) \quad , \tag{13}$$

where m_t stands for the mark-up for wages

$$m_t = \left(1 - \frac{c}{e} \cdot MS_t\right)^{-1} \quad ,$$

where e denotes the price elasticity of the demand for the products of the firm and MS_t denotes the market share of the firm. We also allow the mark-up m_t to depend on the demand factor A_t, because A_t might have an influence on the pricing power of the firm (cf. Layard *et al.*, 1991, Ch. 7 , Martins *et*

al., 1996, and Haskel *et al.*, 1995). Furthermore, we assume that the firm faces quadratic adjustment costs when it tries to adjust its labour force to the target level L_t^*. We assume these costs to be of the form

$$C_t = \sum_{\nu=0}^{\infty} d^{\nu} \left[c_1 E_t (L_{t+\nu} - L_{t+\nu}^*)^2 + c_2 E_t (L_{t+\nu} - L_{t+\nu-1})^2 \right] \tag{14}$$

when discounted to time t. Here $E_t(\cdot)$ denotes the expectation operator, conditional on information available at time t, and d denotes the discount factor.

These premises lead to a dynamic model for labour demand (cf. Layard *et al.*, 1991, Ch. 7, Hamermesh, 1993, Ch. 2 and Nickell, 1986) that can be approximated by a log-linear model of the form

$$\begin{aligned} \log N_t &= \beta_o + \phi_1 \log N_{t-1} + \phi_2 \log N_{t-2} + \beta_1 \log \left(\frac{W_t}{P_t} \right) + \beta_2 \log \left(\frac{W_{t-1}}{P_{t-1}} \right) \\ &\quad + \beta_3 \log K_t + \beta_4 \log MS_t + \beta_5 \log YS_t + \varepsilon_t \quad , \quad \varepsilon_t \sim i.i.d(0, \sigma^2) \quad , \end{aligned} \tag{15}$$

where the labour force L_t has been measured by the number of employees N_t. The reason why we have included the two last explanatory variables in the model is that when the log-linear approximation of (13) has been formed in the spirit of a Taylor expansion, the first partial derivatives of $L(\cdot,\cdot)$ and the value of $L(\cdot,\cdot)$ at the point of expansion will be calculated at points depending on the mark-up term m_t. This is why all factors affecting m_t can also affect the intercept and the coefficients in the log-linear model (15). Here YS_t denotes the aggregate output of the whole branch and is supposed to act as a proxy for the demand factor A_t.

We want to underline that model (15) is supposed to mimic the temporal variations of labour demand of one single firm. When we apply this model form to different firms, it is obvious that the coefficients and the intercept terms are not necessarily the same for all firms. According to the derivation outlined above, there are at least three potential reasons that might cause variation in the coefficients of the log-linear approximation (15). First of all, the conjecture function (12) must depend on the firm's situation on the product market. Thus, the constant c may in fact vary over firms. For the second, the product assortment might be slightly different for each firm thus causing the price elasticity term e to vary over firms. For the third, the relevant discounting factor d might vary over firms, because it presumably depends on the firm's situation on the loan market through the interest rate available to the firm. Differences in production technology may also result in different coefficients, but this will presumably show mainly between the branches, not so much within them. For these reasons we formulated the model for the whole

panel as follows:

$$
\begin{aligned}
n_{it} &= \mu_i + \phi_{i1} n_{i,t-1} + \phi_{i2} n_{i,t-2} + \beta_1 w_{it} + \beta_2 w_{i,t-1} \\
&\quad + \beta_{i3} k_{it} + \beta_4 ms_{it} + \beta_5 ys_t + \alpha_t + \varepsilon_{it} \quad , \\
(\mu_i \quad \phi_{i1} \quad \phi_{i2} \quad \beta_{i3})' \mid S_i &\sim N_4\big((\mu \quad \phi_1 \quad \phi_2 \quad \beta_3)' + BS_i \ , \ \Omega_o \big) \quad , \qquad (16) \\
\varepsilon_{it} &\sim NID(0, \sigma^2) \quad , \quad t = b_i + 2, ..., b_i + T_i - 1 \quad , \\
b_i \geq 1986 \quad , \quad b_i + T_i - 1 \leq 1998 \quad , \quad & S_i = (n_{i,b_i} \quad n_{i,b_i+1})' \quad , \quad i = 1, ..., 260 \ ,
\end{aligned}
$$

where the subscript i refers to firm and small letters have been used to denote the logarithms of the variables N_{it}, K_{it}, MS_{it} and YS_t. The logarithm of real wages have been denoted by w_{it} for simplicity. We did not actually have information on the market shares, so we had to substitute a proxy for them. The proxy variable was formed by the turnovers of each firm, divided by the value of output of the whole branch. A possible time trend, common to all firms, has also been included in the model (16). The values of this common trend have been treated as fixed parameters α_t. Because we had some computational difficulties in handling more than three random parameters simultaneously, we assumed β_1 ,β_2 , β_4 and β_5 to be constants. This assumption was also supported by the data.

In order to get rid of the intercept terms μ_i we analysed our data in the differenced framework

$$
\begin{aligned}
\Delta n_{it} &= \phi_{i1} \Delta n_{i,t-1} + \phi_{i2} \Delta n_{i,t-2} + \beta_1 \Delta w_{it} + \beta_2 \Delta w_{i,t-1} \\
&\quad + \beta_{i3} \Delta k_{it} + \beta_4 \Delta ms_{it} + \beta_5 \Delta ys_t + \alpha_t^* + \Delta \varepsilon_{it} \quad , \\
(\phi_{i1} \quad \phi_{i2} \quad \beta_{i3})' \mid S_i &\sim N_3\big((\phi_1 \quad \phi_2 \quad \beta_3)' + B^* S_i \ , \ \Omega_o^* \big) \quad , \qquad (17) \\
\varepsilon_{it} &\sim NID(0, \sigma^2) \quad , \quad t = b_i + 3, ..., b_i + T_i - 1 \quad , \\
b_i \geq 1986 \quad , \quad b_i + T_i - 1 \leq 1998 \quad , \quad & S_i = (n_{i,b_i} \quad n_{i,b_i+1})' \quad , \quad i = 1, ..., 260 \ ,
\end{aligned}
$$

where we have denoted $\alpha_t^* = \Delta \alpha_t$ for $t = 1989, ..., 1998$. The estimation results are shown in Table 1. The standard errors of the estimates are indicated in parentheses.

Table 1. ML- estimation results for model (17)

PARAMETER	VARIABLE	Model (17) restriction $\Omega_o^* = 0$	Model (17)
ϕ_1	$\Delta n_{i,t-1}$	0.2221 (0.0904)	0.2222 (0.1224)
ϕ_2	$\Delta n_{i,t-2}$	0.03227 (0.0941)	0.04011 (0.1199)
β_1	$\Delta w_{i,t}$	−0.3123 (0.0175)	−0.2773 (0.0168)
β_2	$\Delta w_{i,t-1}$	0.02917 (0.0185)	0.01678 (0.0172)
β_3	$\Delta k_{i,t}$	0.2556 (0.0384)	0.2290 (0.0674)
β_4	$\Delta ms_{i,t}$	0.3537 (0.0141)	0.3607 (0.0147)
β_5	Δys_t	0.3675 (0.0178)	0.3733 (0.0183)
σ^2	ε_{it}	0.03094	0.02205
−2 log Lik		−2090.5	−2357.9

The ML- estimates for the elements of the covariance matrix of the random coefficients in the model (17) were

$$\widehat{\Omega}_o^* = \begin{pmatrix} 0.1892^2 & -0.00508 & -0.00712 \\ -0.00508 & 0.1711^2 & -0.01006 \\ -0.00712 & -0.01006 & 0.1322^2 \end{pmatrix} ,$$

corresponding to the correlation matrix

$$\begin{pmatrix} 1 & -0.1568 & -0.2847 \\ -0.1568 & 1 & -0.4445 \\ -0.2847 & -0.4445 & 1 \end{pmatrix} .$$

The estimated B^*- matrix was

$$\widehat{B}^* = \begin{pmatrix} -0.0763 & (0.1353) & 0.1104 & (0.1365) \\ 0.1436 & (0.1134) & -0.1611 & (0.1150) \\ -0.0574 & (0.0816) & 0.0352 & (0.0825) \end{pmatrix} ,$$

where the estimated standard errors have been indicated in parentheses.

The estimated ARX- structures were stable for every single firm. The most apparent thing in the results of Table 1 is, that there is a tremendous amount of heterogeneity in the panel, because the LR- test statistic for the hypothesis $\Omega_o^* = 0$ would attain the value 267.47. This value is huge since only 6 parameters are eliminated in the hypothesis.

Next, we wanted to test, whether the parameters ϕ_1, ϕ_2, β_1, β_2, β_3 and β_4 vary by branch. After including the interactions between branch and the variables $\Delta n_{i,t-1}$, $\Delta n_{i,t-1}$, $\Delta w_{i,t}$, $\Delta w_{i,t-1}$, $\Delta k_{i,t}$ and $\Delta ms_{i,t}$ in the model (17), the

attained likelihood maximum increased considerably. The value of $-2 \log \text{Lik}$ was -2799.4 , which leads to a LR- test statistic 441.5 corresponding to a p- value smaller than 10^{-10}. So, the dynamic structure of labour demand seems to vary vastly between firms, and the expected values of the firm-specific coefficients seem to vary over branches.

In order to shed some light on the plausible reasons for the differences in the firm-specific coefficients, we first tried to correlate the firm-specific long term wage elasticities to several factors reflecting the firms' situations on the loan market. Among these factors were the financial leverages and the average interest rates of debts of each firm, but we could not find any clear connections between the elasticities and these variables. Because the data set contains no closer information on the product assortments of each firm, we only tried to sort out, whether the firm's situation on the product market has anything to do with its estimated wage elasticity. To this end, we plotted the elasticities against the logarithmic market share proxies $\log MS_{i,b_i}$, but found no connection between the variables. This result can perhaps be taken to indicate that in the Finnish economy, the conjecture functions vary over branches, but do not depend on the relative size of the firm within the branch.

We did, however, find a weak but statistically highly significant positive overall correlation between the firm-specific long-term wage elasticity and the productivity in the firm. The estimated correlation was 0.2274 and the hypothesis of no correlation between the variables corresponded to a p-value 0.0003. This means that firms with high productivity seem to react less to rising wages than do firms with lower productivity.

Table 2. Long-term firm-specific wage elasticities by branch

BRANCH	NUMBER OF FIRMS	AVERAGE LONG-TERM ELASTICITY	RANGE OF THE FIRM-SPECIFIC ELASTICITIES
Food industry	23	−0.717	(−2.095 , −0.487)
Graphical industry	16	−0.568	(−0.773 , −0.416)
Chemical industry	9	−0.231	(−0.317 , −0.192)
Advertising	24	−0.544	(−0.825 , −0.383)
Metal & engineering	32	−0.331	(−0.462 , −0.204)
Forest industry	8	−0.511	(−0.780 , −0.408)
Retail trade	28	−0.484	(−0.622 , −0.400)
Other branches	120	−0.299	(−1.521 , −0.076)
All branches	260	−0.403	(−2.095 , −0.076)

Table 2 reveals that the demand for labour seems to be strongly elastic with respect to real wages especially in food, graphical and forest industries and in advertising, whereas in metal and engineering and in other industries the estimated average elasticity is roughly -0.3 .

CONCLUSION

In this article we have shown that there seem to be individual differences in the labour demand patterns of Finnish firms. A model allowing for firm-specific coefficients outperformed the corresponding model with fixed coefficients by a huge margin. The reason for this heterogeneity needs further investigation, but we did find a correlation between the firm-specific wage elasticity and the productivity of the firm. This is well in line with the results of Honkapohja *et al.* (1999), who found indications that the demand for labour might be less elastic to wages in firms with capital-intensive production than in firms with labour-intensive production.

The model we used in Section 3 can be criticised, because we actually treated the real wages and the gross capital as exogenous variables. However, studying

interdependencies and feedback relations on the basis of heterogeneous panel data would be far more complicated and computationally far more demanding (cf. Rahiala, 2003) than studying unidirectional dependencies.

REFERENCES

Arellano M., Bond S., 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies,* 58, 277-297.

Arellano M., Bond S., 1998, Dynamic Panel Data Estimation Using DPD98 for GAUSS: A Guide for Users. Unpublished manual.

Blundell R., Bond S., 1998, Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* 87, 115-143.

Blundell R., Bond S., 2000, GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews,* 19, 321-340.

Diggle P., Liang K.-Y., Zeger S., 1994, *Analysis of Longitudinal Data.* Oxford University Press, Oxford.

Hamermesh D., 1993, *Labor Demand.* Princeton University Press, Princeton, NJ.

Haskel J., Martin C., Small I., 1995, Price, marginal cost and the business cycle. *Oxford Bulletin of Economics and Statistics,* 57, 25-42.

Henderson H., Searle S., 1981, On deriving the inverse of a sum of matrices. *SIAM Review,* 23, 53-60.

Hoogstrate A., Palm F., Pfann G., 2000, Pooling in dynamic panel-data models: An application to forecasting GDP growth rates. *Journal of Business & Economic Statistics,* 18, 274-283.

Honkapohja S., Koskela E., Uusitalo R., 1999, Työllisyys, työn verotus ja julkisen talouden tasapaino. Valtiovarainministeriö, Tutkimukset ja selvitykset, 1/99.

Layard R., Nickell S., Jackman R., 1991, *Unemployment: Macroeconomic Performance and the Labor Market.* Oxford University Press, Oxford.

Lehman E., 1983, *Theory of Point Estimation.* Wiley, New York.

Martins O., Scarpetta S., Pilat D., 1996, Mark-up pricing, market structure and the business cycle. *OECD Economic Studies,* 26, 71-106.

Nickell S., 1986, Dynamic models of labor demand. In O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics,* vol. 1. North-Holland, Amsterdam.

Pesaran M.H., Smith R., 1995, Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics,* 68, 79-113.

Pesaran M.H., Zhao Z., 1999, Bias reduction in estimating long-run relationships from dynamic heterogeneous panels. In C. Hsiao, K. Lahiri, L.-F. Lee and M. H. Pesaran (eds.), *Analysis of Panels and Limited Dependent Variable Models,* 297-322. Cambridge University Press, Cambridge.

Rahiala M., 1999, Random coefficient autoregressive models for longitudinal data. *Biometrika,* 86, 718-722.

Rahiala M., 2003, Modelling dynamic interdependencies in heterogeneous panel data. In M. Jäntti, G. Rosenqvist and R. Höglund (eds.), *Essays in Honour of Professor Leif Nordberg,* Statistics Finland, Helsinki.

Self S., Liang K.-Y., 1987, Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association,* 82, 605 - 610.

Stram D., Lee J.W., 1994, Variance components testing in the longitudinal mixed effects model. *Biometrics,* 50, 1171-7 ; Correction, 1995, *Biometrics,* 51, 1196.

14. Models with Price Stickiness for Analysis of Monetary Policy

Seppo Honkapohja*

INTRODUCTION

The models that are frequently used in the study of monetary policy changed a great deal during the 1990s. We have seen the development of fully-fledged general equilibrium models in many areas of macroeconomics. Monetary economics is no exception to this development. The current models for the study of monetary policy are typically dynamic general equilibrium models with some forms of market imperfections. Various kinds of market imperfections have been considered in the literature, including contractual limitations in credit markets and monopolistic competition with price stickiness. Recent surveys emphasizing these two types of imperfections are given e.g. in Bernanke, Gertler and Gilchrist (1999), Clarida, Gali and Gertler (1999) and Woodford (2001). This brand of models is often referred to as New Keynesian models, since they focus on role of market imperfections and monetary aspects as a basis of macroeconomic fluctuations.

For brevity, I will focus on the latter type of market imperfection, i.e. price stickiness. My goal is to provide a short exposition of this line of model development. More precisely, my aim is to describe in some detail a typical model of monopolistic competition with sticky prices, so that the theoretical underpinnings of monetary policy can be illuminated.

Starting from first principles, I will develop the foundations of the model. After that I will consider the log-linearized form of the model, which setup is very often the starting point of applied analysis. In these models the form of monetary policy has important consequences for the nature of equilibria in the economy. Using the log-linear model I will take up two basic issues for monetary policy that have been major concerns in the recent literature.

* Seppo Honkapohja is Professor at the University of Helsinki. This paper is written in honour of Pentti Vartia, a colleague and friend for many years, on the occasion of his 60th Birthday.

First, the form of monetary policy affects the number of stationary equilibria in the economy. In particular, under certain settings of monetary policy the stationary equilibrium may not be unique even if the economic fundamentals are stationary. The second important concern is the possible instability of the economy when agents are not fully in equilibrium and try to learn the relevant properties of the economy. Good policy design contributes to the "learnability" of the equilibrium of the economy. Recent research has shown that both of these issues are important for the design of monetary policies, and this is specifically the case when optimal monetary policy and its implementations are considered.

THEORETICAL FOUNDATIONS

A MODEL OF MONEY

Any model of monetary policy must be based on a theory of money, which has been a perennial concern for economic theory. Many deep and less deep theories of money have been developed over the years. The New Keynesian macroeconomic models are not based on any deep theory of money. It is simply assumed that money provides in utility or time savings when carrying out transactions and trading. In macroeconomics this notion has recently been modelled in two ways:

- Money in the utility function (MIU) approach: Real balances are an argument in the utility function of consumers,
- Cash-in-advance (CA) approach: consumers face the constraint requiring that (at least some) consumption goods must be purchased with cash.

Both of these set-ups lead to broadly similar implications, though some details are different. To see the similarity, consider a CA model with two goods, so that one good is a "credit good" and the other is a "cash good". In other words, purchases of the latter must be paid for by money. The instantaneous utility function takes the form $u(C_1, C_2)$ and the cash good must satisfy the cash-in-advance constraint $P_2 C_2 = M$. This constraint allows us to write the utility function as

$$u(C_1, M/P_2),$$

which is close to a standard MIU specification

$$u(C_1, M/P),$$

where P is the price index. In this paper I will use the MIU approach.

The Basic Model

I consider a representative agent economy, so that there is an infinitely lived consumer with a utility function

$$E_t \sum_{s=0}^{\infty} \beta^s u(c_{t+s}, mr_{t+s}),$$

where mr denotes real balances held at the end of the period[1] and c is a consumption aggregate. There is also a production function and each household produces a single good using the function

$$y_t = f(k_t, n_t).$$

Both the utility and production functions are assumed to satisfy the usual boundary (or "Inada") conditions. The model will be simplified by the assumption that capital evolves exogenously and it is not optimised.

The household supplies a fixed quantity of one unit of labor to the market and as producers they purchase labor for production. There is also a government bond on which a real rate of interest r_t is paid, so that its current real price is $(1+r_t)^{-1}$.[2] Inflation is given by $\pi_t = (P_t - P_{t-1})/P_{t-1}$, where P is the nominal price of the good.

The budget constraint of the household is

$$f(k_t, n_t) - tx_t = c_t + k_{t+1} - (1-\delta)k_t + w_t(n_t - 1) + mr_t - mr_{t-1}(1+\pi_t)^{-1} + b_{t+1}(1+r_t)^{-1} - b_t$$

and here b_{t+1}, tx_t denote purchases of bonds in period t and lump-sum taxes paid, respectively.

The optimality conditions are given by the budget constraint and the following first-order conditions:

$$u_1(c_t, mr_t) - \lambda_t = 0$$

$$u_2(c_t, mr_t) - \lambda_t + \beta E_t(\lambda_{t+1}(1+\pi_{t+1})^{-1}) = 0$$

$$-\lambda_t + \beta E_t \lambda_{t+1}\left[f_2(n_{t+1}, k_{t+1}) + 1 - \delta\right] = 0$$

$$f_1(k_t, n_t) - w_t = 0$$

$$\lambda_t + \beta E_t \lambda_{t+1}(1 + r_t) = 0 .$$

Here λ_t are Lagrange multipliers associated with the flow budget constraints.

I will not provide a rigorous justification for the first order conditions, as it requires dynamic programming, see section 2.3 of Chow (1996). Intuitively, one can see how these conditions emerge for the following reason.

Consider the corresponding non-stochastic problem and write a Lagrangian for it:

$$L = \sum_{s=0}^{\infty} \beta^s u(c_{t+s}, mr_{t+s}) + \sum_{s=0}^{\infty} \beta^s \lambda_{t+s}(budj_{t+s}) ,$$

where

$$budj_t = f(k_t, n_t) - tx_t - c_t - k_{t+1} + (1-\delta)k_t - w_t(n_t - 1)$$
$$+ mr_t - mr_{t-1}(1+\pi_t)^{-1} - b_{t+1}(1+r_t)^{-1} + b_t$$

is the flow budget constraint. Optimization in the usual way leads to the first-order conditions above, but without the conditional expectations. To understand why the conditional expectations emerge, note that the feedback control law (and not an open loop policy) is being derived in the stochastic problem. Thus for each period one needs to condition the first-order conditions with conditional information available at that date. This intuitively suggests the placement of the conditional expectations in the conditions above.

One must add the market equilibrium conditions to the first-order conditions above. First, the flow government budget constraint is

$$g_t - tx_t = (1+\pi_t)mr_{t+1} - mr_t + (1+r_t)^{-1} b_{t+1} - b_t .$$

G is the per household government purchases. Second, the market-clearing conditions are

$$y_t = c_t + i_t + g_t, \; n_t = 1, \; mr_t = M_t / P_t .$$

Here M denotes the nominal money supply, and one must also recall the definition of inflation.

It was already mentioned that investment evolves exogenously, i.e. it is assumed to grow at expected constant rate:

$$E_t(\log(i_{t+1}) - \log(i_t)) = \log(1+\xi) \approx \xi .$$

An equilibrium is defined for a given government policy, i.e. specification of taxes, purchases and money supply. If no further assumptions are made about aggregate supply, the model is one of flexible prices. In a moment aggregate supply will be modelled through optimal price setting under costly adjustment. The flexible price version will then yield a linearized model of aggregate demand, to which a model of aggregate supply or nominal price setting can be appended.

The Linearization

Assume that the utility function takes the specific form

$$u(c_t, mr_t) = \sigma(\sigma-1)^{-1}(c_t)^{(\sigma-1)/\sigma} \exp(\omega_t) + (1-\gamma)^{-1}(mr_t)^{1-\gamma},$$

where ω_t is a preference shock. The last first-order condition implies

$$(c_t)^{-1/\sigma} \exp(\omega_t) = \beta E_t (c_{t+1})^{-1/\sigma} \exp(\omega_{t+1})(1+r_t)$$

which yields the log-linearization

$$\omega_t - \sigma^{-1} \log(c_t) = -\sigma^{-1} E_t \log(c_{t+1}) + \log\beta + \log(1+r_t) + E_t\omega_{t+1} .$$

Here approximations of the type $\log E_t d_{t+1} \approx E_t \log d_{t+1}$ are being used. Next, we log-linearize the resource constraint to obtain

$$\log(y_t) = d_1 \log(c_t) + d_2 \log(i_t) + d_3 \log(g_t),$$

which we combine with the linearized Euler equation. If the preference shock is assumed to follow a stationary AR(1) process with parameter $|\rho| < 1$, recalling constant anticipated growth of investment and assuming constant government purchases, we get

$$\log(y_t) = b_0 + b_1 r_t + E_t \log(y_{t+1}) + v_t,$$

where $b_1 = -\sigma d_1$, $v_t = \sigma d_1 (1-\rho)\omega_t$. This equation is often called the *optimising IS curve.*

To obtain the linearized monetary equilibrium condition we take the first, second and last first-order conditions and obtain

$$(mr_t)^{-\gamma} = (c_t)^{-1/\sigma} \exp(\omega_t) - \beta E_t \left[(1+\pi_{t+1})^{-1} (c_{t+1})^{-1/\sigma} \exp(\omega_{t+1}) \right].$$

Then log-linearize, use the definition of the nominal interest rate R_t in terms of the real rate and inflation

$$R_t = r_t + E_t \log(1+\pi_{t+1}) \approx r_t + E_t \pi_{t+1},$$

take note of the Euler equation and use log-linearized market clearing. This gives an expression of the form

$$\log(mr_t) = f_0 + f_1 \log(y_t) - f_2 \log(i_t) - f_3 \log(g_t) + f_3 R_t + f_4 \omega_t .$$

Here f_i are positive constants. (We omit the detailed expressions, which can be explicitly computed by doing the tedious calculations.) This equation is often called the *LM curve.*

Aggregate Supply

The final step is to incorporate price setting or aggregate supply side to the model. There are various formulations in the literature, and see e.g. Roberts (1995) for a few different variants that lead to the same final equation.

One possibility, due to Rotemberg (1982), is to postulate quadratic costs in adjustment of prices, so that firms minimise the costs

$$\min_{\{p\}} \Omega_t = \sum_{s=0}^{\infty} \beta^{t+s} \left[(\log(p_s) - \log(p_s{}^*))^2 + c(\log(p_s) - \log(p_{s-1}))^2 \right] .$$

The first term gives the out-of-equilibrium costs from optimal prices and the second term gives the cost of adjustment. The firm must thus seek a balance between being away from optimum and size of adjustments. The first-order condition is given by

$$E_t \left\{ (\log(p_t) - \log(p_t{}^*)) + c\left[(\log(p_t) - \log(p_{t-1})) - \beta(\log(p_{t+1}) - \log(p_t)) \right] \right\} = 0$$

yielding

$$\Delta \log(p_t) = \beta E_t \Delta \log(p_{t+1}) - (1/c)(\log(p_t) - \log(p_t{}^*)) .$$

If the optimal price is given, say, by a log-linear function

$$\log(p_t^*) = \log(\hat{p}_t) + \alpha \log(y_t) + \varepsilon_t ,$$

where hat denotes the price charged by other firms, *log(y)* is log-deviation of output from aggregate trend and ε_t is *iid* random term, we get (with identical firms) the final equation

$$\Delta \log(p_t) = \beta E_t \Delta \log(p_{t+1}) - (\alpha / c) \log(y_t) - \hat{\varepsilon}_t .$$

This equation is nowadays often called *the new Phillips* or *aggregate supply curve.* It can be noted that the new Phillips curve differs from the Lucas supply function or traditional expectations augmented Phillips curve by having expected future prices mattering for current price changes. This forward-looking aspect comes from the nominal rigidity. The equation also differs from the Lucas formulation by having a coefficient β in front of the expected future price change, where β is the discount factor of the price setter. There is thus a long run trade-off between unemployment and inflation. The trade-off is, however, nearly vertical since in relatively high frequency decisions the discount factor is close to one.

THE LOG-LINEAR MODEL

In summary, the following model has been obtained as an approximation from the preceding analysis. It is written in deviation form, so that constants have been omitted from the equations. The output variable is then the "output gap" and inflation is measured form its trend. The model has three equations: aggregate demand (IS), monetary equilibrium (LM) and aggregate supply (AS). It should be noted that the government budget constraint, in particular its intertemporal form, has been omitted from the model. This is legitimate as long as fiscal policy is passive and set as to ensure the fulfillment of the government budget constraint.

To match most of the literature and to simplify the expressions, we need to revert to new notation. We drop logs from previous notation and introduce the new notation $m_t = \log M_t$, $p_t = \log P_t$ (earlier *mr* referred to real balances). Then we can write the previous equations as

$$y_t = -b(R_t - E_t \pi_{t+1}) + E_t y_{t+1} + u_{1,t}$$

$$m_t - p_t = \tilde{f}_1 y_t - f_2 R_t + u_{2,t} .$$

$$\pi_t = \beta E_t \pi_{t+1} + a y_t + u_{3,t}.$$

The equations can be called the IS curve, the LM curve and the AS curve, respectively. Note that we have included the anticipated signs for the equations, so that all parameters can be taken to be positive. The shocks $u_{i,t}$ are combinations of the random shocks in the previous full model. In applied studies further shocks are also often introduced and the $u_{i,t}$ can also capture these aspects. We now apply this model to discuss some basic aspects of monetary policy.

MONETARY POLICY

The preceding model can be used to study the performance of interest rate rules used in operationally conducting monetary policy. Such rules are typical nowadays, since they have been found to describe how monetary policy is conducted in practice. Naturally, one could equally well study other rules, such as money supply rules. One could also study the performance of optimal monetary policy rules, see Clarida, Gali and Gertler (1999) and Evans and Honkapohja (2002).

An interest rate rule describes how the nominal rate is adjusted in response to, say, output gap or inflation. A simple instrument rule, often called the *Taylor rule* states that nominal interest rate responds to current output gap and inflation[3]

$$R_t = \phi_y y_t + \phi_\pi \pi_t.$$

If monetary policy is conducted by setting the interest rate, then money supply becomes endogenous, and the analysis proceeds by incorporating the interest rate rule to the IS and AS equations. Note that the LM equation is left out, since it just determines the now endogenous nominal money supply that supports this level of the nominal interest rate.

Thus we have the system

$$y_t = -b(\phi_y y_t + \phi_\pi \pi_t - E_t \pi_{t+1}) + E_t y_{t+1} + u_{1,t}$$

$$\pi_t = \beta E_t \pi_{t+1} + a y_t + u_{3,t}.$$

This can be written in matrix form as:

$$\begin{pmatrix} y_t \\ \pi_t \end{pmatrix} = \zeta^{-1} \begin{pmatrix} b^{-1} & 1-\beta\phi_\pi \\ ab^{-1} & a+\beta(b^{-1}+\phi_y) \end{pmatrix} \begin{pmatrix} E_t y_{t+1} \\ E_t \pi_{t+1} \end{pmatrix} + \zeta^{-1} \begin{pmatrix} b^{-1} & -\phi_\pi \\ ab^{-1} & b^{-1}+\phi_y \end{pmatrix} \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}.$$

Here $\zeta = b^{-1} + \phi_y + a\phi_\pi$. From this it is possible e.g. to calculate RE solutions. In numerical version of the model one could easily study responses to shocks etc. Note that in this case, one normally uses a calibrated version of the model and then the precise magnitudes of the parameters in the IS and AS curves need to be obtained. We do not go into empirical analysis here but instead just consider a few basic issues with the model. See e.g. McCallum and Nelson (1999b) or Woodford (1996) for calibrations of the model.

Uniqueness of Equilibria

An important issue for models like the matrix system above is the existence of a unique stationary equilibrium. This issue can be a major concern in models involving expectation since expectations can potentially induce a link from extrinsic phenomena with no fundamental economic significance (often figuratively referred to as 'sunspots') into the determination of endogenous economic variables; here inflation and the output gap. Such extrinsic variables would then influence the equilibrium of the economy solely through expectations, i.e. the equilibrium would depend on a sunspot because agents believe it to matter. The sunspot variable can be interpreted as 'market psychology'. Benhabib and Farmer (1999) provide a survey of the literature on indeterminacy and sunspots.

It can be shown that there exists a unique stationary solution (model is then said to be *determinate*) for some parameter values, but there may be multiplicities (model is then said to be *indeterminate*) for others. Determinacy and indeterminacy depend on the structure of the economy through the model parameters. In this particular model it is important to note that, in particular, the magnitudes of the policy parameters matter for determinacy.

The formal analysis of determinacy and indeterminacy goes as follows. First, we write the matrix form in an abstract way as

$$z_t = JE_t z_{t+1} + Ku_t \quad (*)$$

where

$$z_t = \begin{pmatrix} y_t \\ \pi_t \end{pmatrix},\ u_t = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix},\ J = \zeta^{-1} \begin{pmatrix} b^{-1} & 1-\beta\phi_\pi \\ ab^{-1} & a+\beta(b^{-1}+\phi_y) \end{pmatrix},$$

$$K = \zeta^{-1} \begin{pmatrix} b^{-1} & -\phi_\pi \\ ab^{-1} & b^{-1}+\phi_y \end{pmatrix}.$$

We assume that the shocks are *iid* with mean zero, for simplicity. Then try a solution to (*) of the form

$$z_t = A + Bu_t\ .\ (**)$$

It is easily computed that

$$A = 0,\ B = K,$$

is a unique RE equilibrium of this form when $|I - J| \neq 0$, where I denotes the identity matrix. If there are no other stationary equilibria, then the model and the equilibrium are said to be determinate. This type of equilibrium is called the minimal state variable solution.

A general solution to (*) can be obtained by defining[4]

$$\eta_{t+1} = z_{t+1} - E_t z_{t+1}$$

and writing (*) as

$$z_{t+1} = J^{-1} z_t + \eta_{t+1} - J^{-1} K u_t\ .$$

This is a VAR-type expression and it is non-stationary when both eigenvalues of J^{-1} are outside the unit circle. When this is the case, there is a unique stationary equilibrium that can be obtained by selecting the martingale difference sequence appropriately, so that z_t becomes a linear function of u_t. The easiest way to compute such a fundamental equilibrium is by the method of undetermined coefficients, i.e. guessing that the relation is of the form (**) and verifying the values of the coefficients.

In contrast, if at least one of the eigenvalues lies inside the unit circle, it is possible to find stationary equilibria that are not of the fundamental type. If this is the case, the model is indeterminate.[5]

For the model of monetary policy with Taylor rules, it can be shown that determinacy requires the 'Taylor principle'

$$\phi_\pi > 1,\ \phi_y \geq 0$$

is satisfied. This stipulates that real interest rate is raised in response to inflation and that the policy response of interest rates to output gap is non-negative. This result was shown by Bullard and Mitra (2002).

Stability of Equilibrium under Learning

Finally we consider how to conduct an analysis of learning for this set-up, see the review paper Evans and Honkapohja (2002) for details. The analysis of learning assumes that economic agents behave like econometricians and try to forecast the future values of inflation and the output gap using an econometric model. The parameters of the econometric model are estimated from past data.

The formal analysis can be summarized as follows. Suppose agents have an econometric model of the form (**), but where the coefficients are not equal to the equilibrium values. Agents are assumed to estimate the parameters by running least squares regressions on past data. If agents forecast with this model, then

$$E_t^* z_{t+1} = A,$$

since the best forecast of u_{t+1} is zero as the shock is *iid*. (Note that more complex process for u_{t+1} would lead a more complex form of the forecast function.) Here $E_t^* z_{t+1}$ denotes expectations that may not be rational.

Substituting these forecasts to the model yields the temporary equilibrium of the economy as

$$z_t = JA + Ku_t .$$

Thus we can compute that the mean value for z_{t+1} is JA and the forecast is a rational expectations equilibrium only if $A = 0$.

If expectations are not rational, the agents obtain new values for the parameters A, B by re-estimating the econometric model they use for forecasting. As already noted, it is usually assumed that agents use least squares estimation, which is the classic econometric estimation procedure. This process of learning can be shown to be convergent to the rational expectations equilibrium if the following differential equation

$$\frac{d(A,B)}{d\tau} = (JA,K) - (A,B)$$

is locally asymptotically stable at the equilibrium $A = 0,\ B = K$.[6] The stability condition, known as expectational stability (or E-stability), in this model is that the eigenvalues of the matrix J must have real parts less than one. For the Taylor rule, it can be shown that fulfillment of the Taylor principle leads to stability of the learning process; see Bullard and Mitra (2002).

Optimal Interest Rate Rules

The Taylor rule examined above is just one case of a variety of different interest rate rules that have been examined in the recent literature. Of other possible interest rate policies rules that lead to optimal monetary policy are particularly important.

Taking the viewpoint of stability under learning and determinacy it can be shown that the way that optimal interest setting is implemented is critical. The standard formulation in the literature can run into problems on both counts, i.e. a commonly suggested implementation of optimal policy can often make the economy vulnerable to both indeterminacy and instability under learning. This was demonstrated by Evans and Honkapohja (2003a,b), who also show that there exists, however, a different implementation that delivers always determinacy and learnability of the solution (**) above. This preferred implementation requires that the policy maker reacts to private sector expectations of inflation and the output gap in the right way.

CONCLUSIONS

This paper has provided an overview to recent models of monetary policy. First, I outlined the micro foundations of a typical model used in the literature. Second, I provided an exposition how to analyse the basic issues of stability under learning and determinacy of equilibria in these models. Both issues provide important constraints for the formulation of interest rate setting.

It should be emphasized that this paper is only a brief overview. Many issues have not been discussed. I note here only one important aspect of modelling. The study of learning above presumed that the policy-maker does not need any information about the structural parameters of the economy. Fulfillment of the Taylor principle seemed to suffice and this principle is not dependent of any structural parameters.

However, other interest rate rules, including optimal rules, require knowledge of the structure of the economy and thus, realistically, also the policy-maker (and not only the private agents) faces a learning problem. The

recent advances in the modelling of learning make it possible to formally analyse also the consequences of policy-maker's learning about the structure of the economy; for formal discussion see Evans and Honkapohja (2002). It can be shown that simultaneous learning by the private agents and the policy-maker will lead to convergence of the economy to a full intertemporal equilibrium if the interest rate rule is designed in accordance with the results obtained from the simpler setting where only private agents are learning.

Another possible complication can arise from limited observability of the key variables for policy making. In particular, the policy maker relying on a forecast-based interest rate rule should ideally use private sector forecasts in the rule since it is these forecasts that matter for the behavior of the private agents. However, private forecasts may not be accurately observable. One possible solution would be for the policy-maker to replace private forecasts by its own internal forecasts. Honkapohja and Mitra (2003) have shown that this procedure can very often lead to convergence of the economy if the interest rate setting respects the guidelines from simpler models.

To conclude, I want to suggest that the new results and approaches to the study of monetary policy that were reviewed in this paper are promising. It should be noted that they have thus far been derived in simple and stylized theoretical models. Much further work is required before it is possible to assess with any confidence how useful these new approaches are as framework and guidance to practical policy-making.

REFERENCES

Benhabib, J. and Farmer, R. (1999): Indeterminacy and Sunspots in Macroeconomics, in Taylor, J.B. and M. Woodford (eds.), Handbook of Macroeconomics, vol. IA, Elsevier, 387-448.

Bernanke, B.S., Gertler, M. and Gilchrist, S. (1999): The Financial Accelerator in a Quantitative Business Cycle Framework, in Taylor, J.B. and M. Woodford (eds.), Handbook of Macroeconomics, vol. IC, Elsevier, 1341-1393.

Bullard, J. and Mitra, K. (2002): Learning about Monetary Policy Rules, Journal of Monetary Economics, 49, 1105-1129.

Chow, G. (1996): Dynamic Economics, Oxford University Press.

Clarida G., Gali, J. and Gertler, M. (1999): The Science of Monetary Policy: A New Keynesian Perspective, Journal of Economic Literature, 37, 1661-1707.

Evans, G.W. and Honkapohja, S. (2001): Learning and Expectations in Macroeconomics, Princeton University Press.

Evans, G.W. and Honkapohja, S. (2002): Adaptive Learning and Monetary Policy Design, paper for the 2002 Journal of Money, Credit and Banking Conference, available in *www.valt.helsinki.fi/raka/seppo.htm*.

Evans, G.W. and Honkapohja, S. (2003a): Expectations and the Stability Problem for Optimal Monetary Policy, Review of Economic Studies, forthcoming.

Evans, G.W. and Honkapohja, S. (2003b): Monetary Policy, Expectations and Commitment, mimeo, available in *www.valt.helsinki.fi/raka/seppo.htm*.

Honkapohja, S. and Mitra, K. (2003): Performance of Monetary Policy with Internal Central Bank Forecasting, mimeo, available in *www.valt.helsinki.fi/raka/seppo.htm*.

McCallum, B.T. and Nelson, E. (1999a): An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis, Journal of Money, Credit and Banking, 31, 296-316.

McCallum, B.T. and Nelson, E. (1999b): Performance of Operational Policy Rules in an Estimated Semi-classical Model, in Taylor J.B. (ed.): Monetary Policy Rules, University of Chicago Press. Also available as NBER Working Paper Nr.6599, June 1998.

Roberts, J.M. (1995): New Keynesian Economics and the Phillips Curve, Journal of Money, Credit and Banking, 27, 975-984.

Rotemberg, J. (1982): Sticky Prices in the United States, Journal of Political Economy, 60, 1187-1211.

Woodford, M. (1996): Control of the Public Debt: A Requirement for Price Stability?, NBER Working Paper Nr.5684.

Woodford, M. (2001): Optimizing Models with Nominal Rigidities, available in *www.princeton.edu/~woodford/*.

ENDNOTES

[1] We note that an alternative is to use beginning-of-period balances as the argument in the utility function. Both assumptions are common in the literature. Subtle differences emerge from the implications, compare the two papers by McCallum and Nelson (1999a, b) in the references.

[2] We follow the McCallum and Nelson (199a) here with the indexing of the return.

[3] Note that nominal interest pegging is a special case of this rule.

[4] η_{t+1} is then a martingale difference sequence, but otherwise arbitrary.

[5] A systematic treatment is beyond the scope of the course. See e.g. Evans and Honkapohja (2001), chapter 10, appendix 2 for a full treatment of determinacy and indeterminacy.

[6] Here τ refers to notional time. The techniques for the study of least squares and related learning schemes are explained in detail in Evans and Honkapohja (2001).

15. Tuloerot ja tulonjako alueittain: Jakautuiko Suomi 1990-luvulla?

Heikki A. Loikkanen ja Risto Sullström*

JOHDANTO

Alueellinen kasautuminen kaupunkialueilla sijaitsevien työpaikkojen ja väestön osuuden kasvun muodossa on ollut kaikissa maissa taloudelliseen kasvuun liittyvä rakennepiirre. 1900-luvulla henkeä kohti lasketun bruttokansantuotteen kasvun rinnalla urbanisoituminen on edennyt rauhan aikana lähes keskeytyksettä koko viime vuosisadan ajan. Kaupunkiväestön osuus on keskimäärin sitä suurempi mitä korkeamman tulotason maasta on kyse.[1]

Suomessa kaupunkiväestön osuus on ollut taloudelliseen kehitysasteeseen nähden kansainvälisesti vertaillen alhainen, mutta urbanisoitumisen perusmalli on ollut meillä sama kuin muuallakin: nousukausina muuttoliike urbaaneille alueille kiihtyy, kun taas taantumissa prosessi hidastuu tai pysähtyy väliaikaisesti kokonaan. Tämän perusmallin mukaisesti 1990-luvun laman aikana muuttoliike pysähtyi, ja sen jälkeinen talouden nousukausi johti kasautuvaan aluekehitykseen. Kaupungistumisen aaltojen suhdanneriippuvuus näyttää jatkuvasti hämärtäneen Suomessa sen mitä on tapahtumassa. Muuttoliikkeen voimistuminen tulee aina yllätyksenä, johon ei ole varauduttu. Toisaalta viime vuosikymmenen alun laman syvyys ja vastaavasti sitä seuranneen korkeasuhdanteen voimakkuus ja kesto olivat poikkeuksellisia. Uutta oli myös se, että muuttoliike suuntautui 1990-luvun lopulla vain puoleen tusinaan kasvukeskukseen eikä useampiin kaupunkialueisiin kuten aiemmissa korkeasuhdanteissa oli tapahtunut. Tässä mielessä maan voi sanoa jakautuneen entistä selvemmin ja osin uudella tavalla voittajiin ja enemmän tai vähemmän hävinneisiin.

Kaupungistumista vähemmän tunnettu pitkän aikavälin kehitysilmiö taloudellisesti kehittyneiden maiden ja niiden alueiden kesken on ollut alueellisten tuloerojen kaventuminen. Myös Suomessa henkeä kohti lasketun

* Heikki A. Loikkanen on professori Helsingin yliopistossa ja Risto Sullström erikoistutkija Valtion taloudellisessa tutkimuskeskuksessa (VATT).

BKT:n, verotulojen ja käytettävissä olevien tulojen alueelliset erot ovat kaventuneet II maailmansodan jälkeisenä aikana. Mutta mitä 1990-luvun lamassa ja sen jälkeisessä nousussa tapahtui alueellisille tuloeroille? Jakautuiko maa tässä suhteessa jollain tavoin?

Tässä kirjoituksessa tarkastellaan käytettävissä olevien tulojen avulla alueiden välisiä suhteellisia tuloeroja ja alueiden sisäisiä suhteellisia tuloeroja 1990-luvulla eli talouden lamassa ja sen jälkeisessä nousussa aina vuosituhannen loppuun saakka.[2] Aluejakoina käytetään sekä pohjois-eteläsuunnassa maan jakavaa luokitusta että kaupungit, taajamat ja maaseutu -kolmijakoa. Tarkoituksena on kartoittaa, johtiko 1990-luvun lama-aika tai sen jälkeinen nousukausi alue-erojen kasvuun tai alueiden sisäisten tuloerojen kasvuun. Tulonjakotarkastelua täydennetään köyhien ja rikkaiden lukumäärien ja väestöosuuksien tarkastelulla.

AINEISTO, ALUEJAOT JA SOVELLETTAVAT KÄSITTEET

Vuosittaisia kotitaloustason tietoihin perustuvia tulovertailuja voidaan tehdä otospohjaisten kulutusaineistojen ja tulonjakotilastoaineistojen perusteella. Edellisiä on laadittu noin viiden vuoden välein vuodesta 1966 alkaen ja jälkimmäisiä vuodesta 1977 lähtien vuosia 1981 ja 1985 lukuun ottamatta. Tässä artikkelissa hyödynnetään ainoastaan vuosien 1990-2000 tulonjakotilastoja, koska tarkoituksena on keskittyä viime vuosikymmenen kehitykseen.[3]

Tulonjakotilastot perustuvat otosaineistoihin. Niiden vuosittainen koko on tarkasteluajanjakson aikana vaihdellut 8 964-11 749 kotitalouden välillä vuosina 1990-2000. Otoskoko rajoittaa osaltaan alueiden määrä ja kokoa, joihin maa voidaan jakaa. Tässä artikkelissa käytetään rinnakkain kahta aluejakoa, suuraluejakoa ja tilastollista kuntaryhmitystä (Tilastokeskus, 1998).

Suuraluejako jakaa Suomen maantieteellisesti neljään alueeseen pohjois-eteläsuunnassa:[4]

1. Pääkaupunkiseutuun
2. Muuhun Etelä-Suomeen
3. Väli-Suomeen
4. Pohjois-Suomeen

Toisena ryhmittelynä sovelletaan tilastollista kuntaryhmitystä, joka jakaa maan kunnat niiden taajama-asteen mukaan:[5]

1. Kaupunkimaisiin kuntiin
2. Taajaan asuttuihin kuntiin
3. Maaseutumaisiin kuntiin

Kotitalouksien taloudellisen aseman arvioinnissa sovelletaan vain yhtä tulokäsitettä eli käytettävissä olevia tuloja. Niitä ei deflatoitu alueellisilla hintaindekseillä, koska sellaisia ei ole käytettävissä. Tämä tulokäsite pitää sisällään myös julkisen sektorin keskeisten toimenpiteiden eli välittömän verotuksen ja tulonsiirtojen vaikutukset, sillä määritelmällisesti

kotitalouden käytettävissä olevat tulot =
tuotannontekijätulot + tulonsiirrot – välittömät verot.[6]

Huomattakoon, että välillisen verotuksen ja saatujen subventoitujen (tai maksuttomien) hyvinvointipalvelujen vaikutukset eivät tule esiin keskityttäessä käytettävissä olevien tulojen tarkasteluun.

Käytettävissä olevia tuloja voidaan tarkastella esimerkiksi kotitaloutta tai henkilöä kohti. Edellinen jättää perhekoon huomiotta kokonaan ja jälkimmäinen puolestaan ottaa sen huomioon, mutta olettaa, että kotitalouksien rahankäytössä ei esiinny lainkaan skaalaetuja. Tässä artikkelissa sovelletaan perinteistä OECD:n ekvivalenssiskaalaa, jolla pyritään saamaan erilaiset kotitaloudet kulutustarpeen suhteen keskenään verrattaviksi. Kotitalouden käytettävissä olevat tulot jaetaan kulutusyksikköjen määrällä, joka saadaan laskemalla yhteen seuraavat painot: ensimmäinen aikuinen saa painon yksi, kaikki seuraavat aikuiset 0,7 ja kukin 0-17 -vuotias lapsi painon 0,5. Näin esimerkiksi kahden aikuisen ja kahden lapsen kotitalouden tapauksessa tulot jaetaan 2,7:llä. Tämä kulutusyksiköllä muunnettu tulo tulee kaikkien kotitalouden jäsenten eduksi ja painotetaan laskelmissa henkilö- ja otospainoilla.

Alueiden välisiä tuloeroja mitattaessa keskitytään suhteellisten tuloerojen tarkasteluun. Tällöin verrataan kunkin aluekokonaisuuden per capita lasketun käytettävissä olevan tulon suhdetta maan keskiarvoon ja analysoidaan sitä, ovatko alueiden erot kasvaneet vai supistuneet 1990-luvun aikana. Tulonjaon suhteellista epätasaisuutta eli eriarvoisuutta mitataan Ginikertoimen avulla, joka saa arvoja välillä 0 ja 1. Täydellisen tasajaon tapauksessa kerroin saa arvon 0 ja täydellisen eriarvoisuuden tapauksessa arvon 1.

Tulonjakotarkastelua täydennetään analysoimalla köyhien ja rikkaiden henkilöiden lukumäärän kehitystä 1990-luvulla. Tässä yhteydessä köyhäksi määritellään henkilöt, joiden käytettävissä olevat tulot per kulutusyksikkö ovat alle köyhyysrajan, joka on puolet koko maan mediaanitulosta vuonna 1990 (vrt. Sullström, 1989). Rikkaaksi määritellään henkilöt, joiden käytettävissä olevat tulot per kulutusyksikkö ovat yli kaksi kertaa koko maan mediaanitulon vuonna 1990. Kiinnittämällä köyhyyden ja rikkauden rajat vuoden 1990 reaaliseen tulotasoon mediaanitulon muutokset muina vuosina eivät muuta ko. rajoja.

ALUEIDEN VÄLISET TULOEROT

Vuosi 1990 oli lamaa edeltänyt viimeinen vuosi, jonka aikana kokonaistuotanto (BKT) vielä kasvoi Suomessa. Sen jälkeisinä kolmena vuotena BKT pieneni kääntyen nopeaan nousuun vuodesta 1994 alkaen. Myös kotitalouksien käytettävissä olevat tulot laskivat 1990-luvun laman aikana ja kääntyivät sittemmin nousuun ripeän taloudellisen kasvun myötä. Kasvoivatko alueelliset tuloerot tänä aikana merkittävästi? Jakautuiko maa kahtia tai useampaan osaan siinä mielessä että alue-erot repesivät?

Pohjois-eteläsuunnassa katsoen pääkaupunkiseudulla tulotaso (per capita) oli vuonna 1990 noin 25 prosenttia maan keskiarvon yläpuolella (taulukko 15.1). Syvimpänä lamavuotena 1993 sen suhteellinen asema oli sama ja heikkeni vielä ripeän taloudellisen kasvun alkuvaiheessa ollen 17 prosenttia maan keskiarvon yläpuolella vuonna 1995. Vasta tarkastelujakson loppupuolella tulotason nousu kiihtyi, ja vuonna 2000 oltiin jo 28 prosenttia maan keskiarvon yläpuolella.

Heikoimmassa suhteellisessa asemassa vuonna 1990 oli Pohjois-Suomi, jonka per capita tulotaso oli 10 prosenttia maan keskiarvon alapuolella. Tämä asema säilyi yli lama-ajan alkaen heiketä vuosikymmenen lopussa päätyen yli 13 prosenttia maan keskiarvon alapuolelle. Väli-Suomi päätyi vuonna 2000 samaan suhteelliseen asemaan, jossa se oli kymmenen vuotta aikaisemminkin. Muu Etelä-Suomi pyöri maan keskiarvon ympärillä kunnes jäi hieman sen alle jakson lopussa.

Taulukko 15.1. Käytettävissä olevat tulot (per capita) suuralue- ja taajama-astejaolla vuosina 1990, 1993, 1995, 1997 ja 2000, (Suomi = 100)

	1990	1993	1995	1997	2000
Suuraluejako					
Pääkaupunkiseutu	125,7	125,3	117,0	119,2	128,4
Muu Etelä-Suomi	98,9	99,5	100,3	100,1	97,2
Väli-Suomi	90,8	89,8	93,2	92,0	90,6
Pohjois-Suomi	89,5	89,3	89,2	88,4	86,6
Taajama-astejako					
Kaupunkimaiset kunnat	107,4	107,2	106,4	106,1	107,3
Taajaan asutetut kunnat	95,1	93,0	95,7	91,1	90,1
Maaseutumaiset kunnat	87,9	88,9	89,0	90,5	88,1
Suomi	100,0	100,0	100,0	100,0	100,0

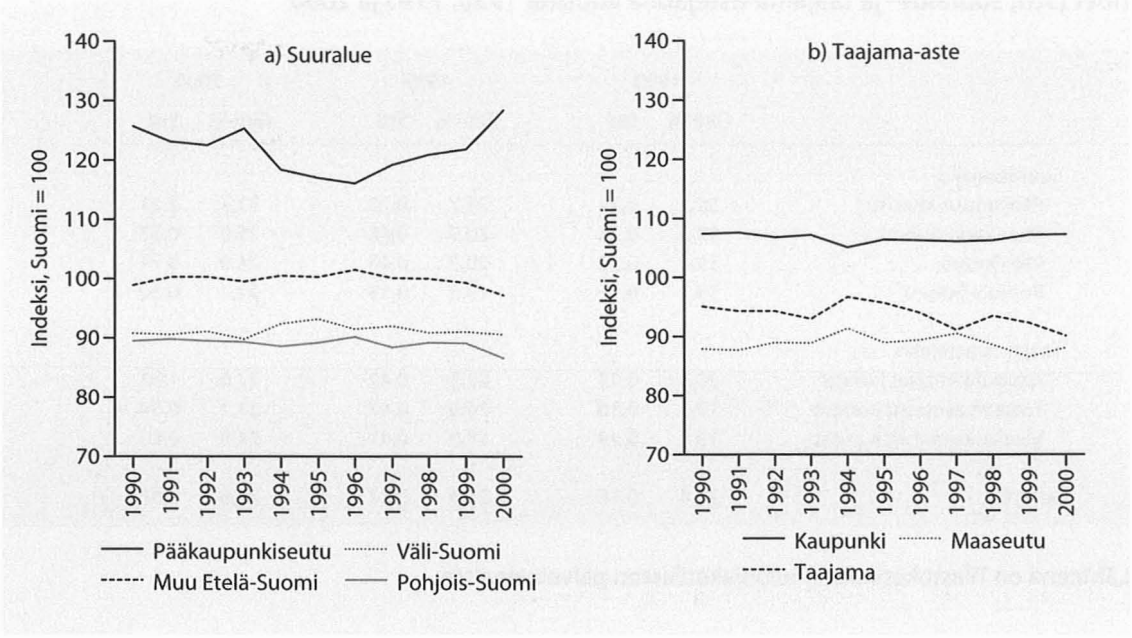
Lähteenä on Tilastokeskuksen tulonjakotilaston palveluaineisto.

Pohjois-eteläsuunnassa katsoen lama-aika ei jakanut Suomea kahtia tai useampiin selvästi eri suuntiin tulotasoltaan kehittyviin osiin. Vasta 1990-luvun lopun nousukauden lopulla voi sanoa pääkaupunkiseudun ja Pohjois-Suomen välisen suhteellisen eron kasvaneen siitä, mitä se oli vuonna 1990 eli 36,2 prosenttiyksiköstä 41,8 prosenttiyksikköön.

Kaupunkien, taajamien ja maaseudun suhteelliset tuloerot 1990-luvulla ilmenevät myös taulukosta 15.1. Kaupunkimaisten kuntien ryhmän suhteellinen asema säilyi lähes muuttumattomana eli tulotaso oli 6-7 prosenttia maan keskiarvon yläpuolella kaikkina tarkasteltuina vuosina. Maaseutumaisten kuntien asema säilyi likipitäen yhtä vakaana, 10-12 prosenttia maan keskiarvon alapuolella koko tarkasteluajan. Sitä vastoin taajaan asuttujen kuntien ryhmässä suhteellinen tuloasema heikkeni jo laman aikana, palautui lähtötasolleen ja sitten heikkeni selvästi noususuhdanteessa, jääden noin 10 prosenttia maan keskiarvosta vuonna 2000.

Kuviossa 15.1 suhteellisten tuloerojen kehitys esitetään molemmilla aluejaoilla aikasarjoina, jotka kattavat kaikki vuodet 1990-2000. Sen perusteella on tehtävissä johtopäätös, jonka mukaan Suomi ei per capita tulojen perusteella jakautunut 1990-luvun lamassa pohjois-eteläsuunnassa, eivätkä erot kaupunkien ja maaseudun välillä kärjistyneet tänä aikana. Vasta vuosituhannen viimeisinä vuosina ei laman, vaan korkeasuhdanteen myötä alue-erot

Kuvio 15.1. Käytettävissä olevat tulot (per capita) suuralue- ja taajama-astejaolla vuosina 1990-2000 (Suomi = 100)



Lähteenä on Tilastokeskuksen tulonjakotilaston palveluaineisto.

kasvoivat niin, että etenkin pääkaupunkiseudun ja Pohjois-Suomen ero revähti selvästi vuoden 1990 asetelmaa suuremmaksi. Kaupunkien ja maaseudun suhteellisessa asemassa ei tapahtunut koko vuosikymmenellä juuri mitään. Sitä vastoin taajamaväestön suhteellinen asema, joka hieman parani 1990-luvun puolivälissä, on sen jälkeen heikentynyt ollen vuonna 2000 samassa asemassa maaseutuväestön kanssa.

TULONJAKO KOKO MAASSA JA ALUEITTAIN

Tulonjakoa tarkastellaan sekä koko maan tasolla että alueittain. Tulonjaon epätasaisuuden mittana käytetään Gini-kerrointa (esim. Lerman ja Yitzhaki 1984), jonka arvot ilmaistaan prosenttimuodossa. Taulukossa 15.2 tulonjakoa arvioidaan 1990-luvulla viiden vuoden välein.

Koko maan tasolla Gini-kertoimen arvo ei yllättäen kasvanut juuri lainkaan siirryttäessä vuodesta 1990 vuosikymmenen puoliväliin, vaikka samaan aikaan työttömyys kasvoi yli 15 prosentin tasolle alle viiden prosentin lukemista vuonna 1990. Kaikkien tuloryhmien käytettävissä olevissa tuloissa tapahtui laskua ja toisaalta sosiaaliturvajärjestelmä ylläpiti työttömien tulotasoa niin, että suhteelliset tuloerot ja Gini-kerroin säilyivät liki ennallaan. Lama-ajan jälkeen alkoi selvä tuloerojen kasvu, ja Gini-kerroin koko maassa nousi viidellä prosenttiyksiköllä vuodesta 1995 vuoteen 2000.

Taulukko 15.2. Käytettävissä olevien tulojen Gini-kertoimet (%) ja asymptoottiset keskivirheet (Std) suuralue- ja taajama-astejaolla vuosina 1990, 1995 ja 2000

	1990		1995		2000	
	Gini-%	Std	Gini-%	Std	Gini-%	Std
Suuraluejako						
Pääkaupunkiseutu	20,2	0,34	23,7	0,70	31,7	2,21
Muu Etelä-Suomi	19,3	0,24	20,9	0,48	25,0	0,57
Väli-Suomi	19,4	0,29	20,7	0,47	24,0	0,71
Pohjois-Suomi	19,7	0,51	19,5	0,53	22,7	0,60
Taajama-astejako						
Kaupunkimaiset kunnat	20,3	0,23	22,3	0,42	27,6	1,01
Taajaan asutetut kunnat	19,3	0,36	20,0	0,67	23,2	0,54
Maaseutumaiset kunnat	19,8	0,29	19,8	0,41	24,9	0,80
Suomi	20,4	0,18	21,6	0,27	26,6	0,67

Lähteenä on Tilastokeskuksen tulonjakotilaston palveluaineisto.

Aluetarkastelussa ensimmäinen huomion arvoinen seikka on se, että vuonna 1990 sekä suuraluejaolla että taajamajaolla tarkastellen alueiden sisäisissä tuloeroissa ei ollut juuri lainkaan eroja. Kaikki erot koko maan Gini-kertoimeen olivat suurimmillaankin vain yhden prosenttiyksikön suuruisia.

Suuraluejaolla tarkastellen vuoteen 1995 mennessä eriarvoisuus alueiden sisällä kasvoi vain lievästi, mutta selvä Gini-kertoimen nousu tapahtui talouden nousun myötä. Vuonna 2000 erityisesti pääkaupunkiseudun sisäinen eriarvoisuus oli yli 10 prosenttiyksikköä vuoden 1990 lähtötason yläpuolella, kun taas muiden alueiden vastaavat nousut olivat 3-6 prosenttiyksikön välillä.

Taajama-asteen mukaisessa tarkastelussa maaseutumaisissa kunnissa ei tapahtunut mitään muutosta eriarvoisuudessa 1990-luvun alkupuolella ja Gini-kertoimien muutokset kaupunkialueilla (lisäystä kaksi prosenttiyksikköä) ja taajaan asutuissa kunnissakin (alle prosenttiyksikön) olivat pieniä. Vuosituhannen lopulla tilanne oli muuttunut tälläkin jaolla. Kaupunkien sisäinen eriarvoisuus kasvoi vuoteen 2000 noin seitsemän prosenttiyksikköä vuodesta 1990. Taajamissa vastaava muutos oli neljä ja maaseutumaisissa kunnissa noin viisi prosenttiyksikköä.

Kuviosta 15.2 sivulla 280 ilmenevät Gini-kertoimien vuosittaisiin arvoihin perustuvat aikasarjat. Koko maan aikasarja osoittaa, että laman alkuvaiheessa eriarvoisuus ei kasvanut lainkaan. Kun tuloerot vuonna 1993 alkoivat kasvaa, suurin muutos tapahtui pääkaupunkiseudulla. Muut suuralueet kokivat samansuuntaisen kehityksen vasta talouden käännyttyä voimakkaaseen kasvuun pari vuotta myöhemmin. Taajama-astejaolla kaupunkialueiden eriarvoisuus oli koko 1990-luvun ajan maaseutua ja taajaan asuttuja kuntia suurempaa. Molemmilla aluejaoilla ääripäiden erot eli pääkaupunkiseudun ero Pohjois-Suomen ja toisaalta kaupunkien ero taajamiin olivat suurimmillaan ajanjakson lopussa.

KÖYHÄT JA RIKKAAT SUURALUEITTAIN

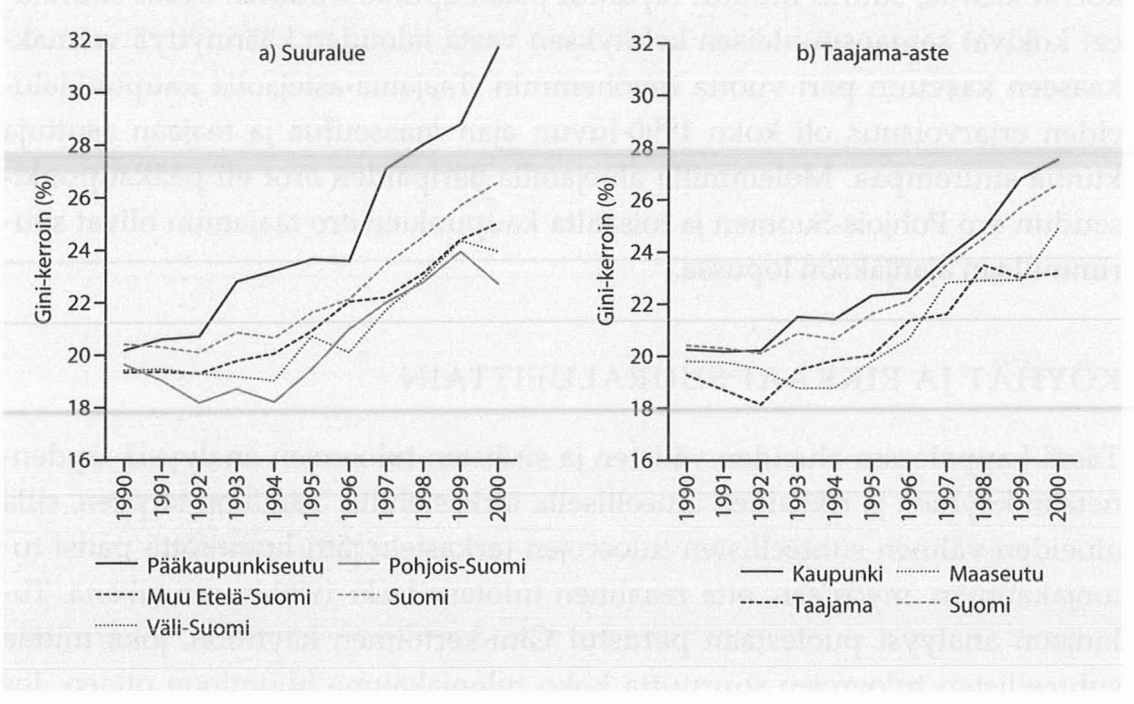
Tässä kappaleessa alueiden välisten ja sisäisten tuloerojen analyysiä täydennetään köyhien ja rikkaiden alueellisella tarkastelulla. Tämä on tarpeen, sillä alueiden välinen suhteellisten tuloerojen tarkastelu jätti huomiotta paitsi tulonjakauman, myös sen, että reaalinen tulotaso laski 1990-luvun aikana. Tulonjaon analyysi puolestaan perustui Gini-kertoimen käyttöön, joka mittaa suhteellisten tuloerojen suuruutta koko tulonjakauma huomioon ottaen. Jos kaikkien tulot laskevat, kuten laman alkuvaiheessa tapahtui, Gini-kertoimessa ei tapahdu selvää muutosta. Määrittelemällä köyhyysraja (puolet mediaa-

nitulosta) ja rikkausraja (kaksi kertaa mediaanitulo) vuoden 1990 tilanteessa ja soveltamalla vastaavia (reaalisia) rajoja läpi koko tarkasteluajanjakson, saadaan täydentävä kuva siitä miten lama-aika ja sen jälkeinen nousu vaikuttivat huonossa ja hyvässä taloudellisessa tilanteessa olevien määriin ja osuuksiin viime vuosikymmenellä (Riihelä ja Sullström, 2001; Riihelä, Sullström ja Tuomala 2001).

Taulukossa 15.3 esitetään köyhien ja rikkaiden lukumääriä koskevat tiedot kolmena ajankohtana. Vuonna 1990 köyhiä oli noin 126 000 henkilöä. Vuonna 1995 köyhien määrä oli reilut 40 000 ja vuonna 2000 vajaat 20 000 henkilöä suurempi kuin vuonna 1990. Jos vuoden 1990 reaalisia käytettävissä olevia tuloja merkitään 100:lla, niin vuonna 1995 indeksi oli 96 ja vuonna 2000 jo 115. Lamasta johtuva reaalitulojen lasku näkyi köyhien lukumäärän kasvuna vuonna 1995. Yllättävämpää on, että vuonna 2000 ripeästä reaalitulojen kasvusta huolimatta köyhien lukumäärä ei vielä laskenut vuoden 1990 tasolle.

Suuralueittain köyhien määrä kasvoi vuodesta 1990 vuoteen 1995 kaikilla alueilla erityisesti pääkaupunkiseudulla ja Pohjois-Suomessa. Sitä vastoin kaikilla muilla suuralueilla paitsi Väli-Suomessa köyhien määrät ja osuu-

Kuvio 15.2. Käytettävissä olevien tulojen Gini-kertoimet (%) suuralue- ja taajama-astejaolla vuosina 1990-2000



Lähteenä on Tilastokeskuksen tulonjakotilaston palveluaineisto.

det laskivat verrataessa vuosia 1995 ja 2000. Taajama-astejaolla taulukon 15.3 numerotietoja tarkastellen köyhien määrä kasvoi vuodesta 1990 vuoteen 1995 kaikilla alueilla. Talouden nousun aikana köyhien lukumäärän kasvu koski vain taajaan asutettuja alueita.

Rikkaita oli vuonna 1990 noin 156 000, ja määrä odotetusti laski noin 143 000 henkilöön vuonna 1995. Sen jälkeen kasvu kiihtyi niin, että vuonna 2000 rikkaita oli jo 352 000. Kymmenessä vuodessa rikkaiden osuus kasvoi kolmesta liki seitsemään prosenttiin. Suuralueittain rikkaita oli jo vuonna 1990 suhteellisesti eniten pääkaupunkiseudulla (liki 8 prosenttia). Vuonna 1995 pääkaupunkiseudun rikkaiden määrässä ja väestöosuudessa tapahtui lievää vähenemistä. Talouden nousun lopulla rikkaiden osuus kasvoi yli 14 prosenttiin vuonna 2000. Myös muilla suuralueilla muutos vuodesta 1990 vuoteen 1995 oli pieni ja pääosin lievästi aleneva, mutta sitten rikkaiden määrä ja osuus kasvoivat. Muun Etelä-Suomen ja Väli-Suomen rikkaiden määrä 2,5-kertaistui seuraavan viiden vuoden aikana ja Pohjois-Suomessa määrä lähes kolminkertaistui.[7]

Taulukko 15.3. Köyhien ja rikkaiden lukumäärät ja väestöosuudet suuralue- ja taajama-astejaolla vuosina 1990, 1995 ja 2000

	1990		1995		2000	
	Henkilöä	%	Henkilöä	%	Henkilöä	%
Köyhät*						
Pääkaupunkiseutu	9 744	1,2	22 262	2,5	15 720	1,7
Muu Etelä-Suomi	57 810	2,6	70 836	3,2	60 516	2,7
Väli-Suomi	42 505	3,2	44 879	3,4	51 034	4,0
Pohjois-Suomi	16 407	2,6	30 595	5,0	15 728	2,5
Kaupunkimaiset kunnat	57 062	2,0	92 931	3,2	74 956	2,4
Taajaan asutetut kunnat	16 207	2,0	19 549	2,4	28 880	3,4
Maaseutumaiset kunnat	53 197	3,9	56 092	4,1	39 153	3,3
Suomi	126 466	2,5	168 572	3,3	142 989	2,8
Rikkaat**						
Pääkaupunkiseutu	64 837	7,8	57 537	6,5	134 616	14,3
Muu Etelä-Suomi	55 758	2,5	53 346	2,4	136 598	6,1
Väli-Suomi	23 091	1,7	23 895	1,8	56 584	4,4
Pohjois-Suomi	12 936	2,0	8 533	1,4	24 425	3,9
Kaupunkimaiset kunnat	113 362	4,1	106 660	3,7	261 099	8,5
Taajaan asutetut kunnat	18 130	2,2	15 681	1,9	35 708	4,2
Maaseutumaiset kunnat	25 130	1,8	20 970	1,5	55 416	4,7
Suomi	156 622	3,1	143 311	2,8	352 223	6,9

Lähteenä on Tilastokeskuksen tulonjakotilaston palveluaineisto. * Köyhyysraja: käytettävissä olevat tulot yhtä suuret kuin puolet mediaanitulosta vuonna 1990. ** Rikkausraja: käytettävissä olevat tulot yhtä suuret kuin kaksi kertaa mediaanitulo vuonna 1990.

Taajama-asteen mukaan tarkastellen rikkaita on lukumääräisesti kaikkina vuosina eniten kaupungeissa, noin 113 000 vuonna 1990 ja 261 000 vuonna 2000. Rikkaiden määrät ja osuudet kasvoivat vuosikymmenen loppupuoliskolla selvästi myös taajaan asutetuissa kunnissa ja maaseudulla.

Taulukon 15.3 lukujen perusteella näyttää siis käyneen niin, että 1990-luvun puoliväliin mennessä köyhien lukumäärä hienokseltaan kasvoi ja rikkaiden määrä puolestaan laski. Vuosikymmen loppupuolella köyhien määrä pysyi lähes ennallaan tai laski ja rikkaiden määrä puolestaan kasvoi.

Kuviossa 15.3 on vuosittaisiin tietoihin perustuvat aikasarjat köyhien lukumääristä suuraluejaolla ja taajama-astejaolla. Vasemmanpuoliset kuviot osoittavat, että köyhyys ei työttömyyden tavoin räjähtänyt kasvuun 1990-luvun laman aikana, vaikka selvää nousua vuosina 1990-1996 onkin havaittavissa yhtäältä pääkaupunkiseudun ja Pohjois-Suomen sekä toisaalta kaupunkimaisten kuntien köyhien lukumäärissä.

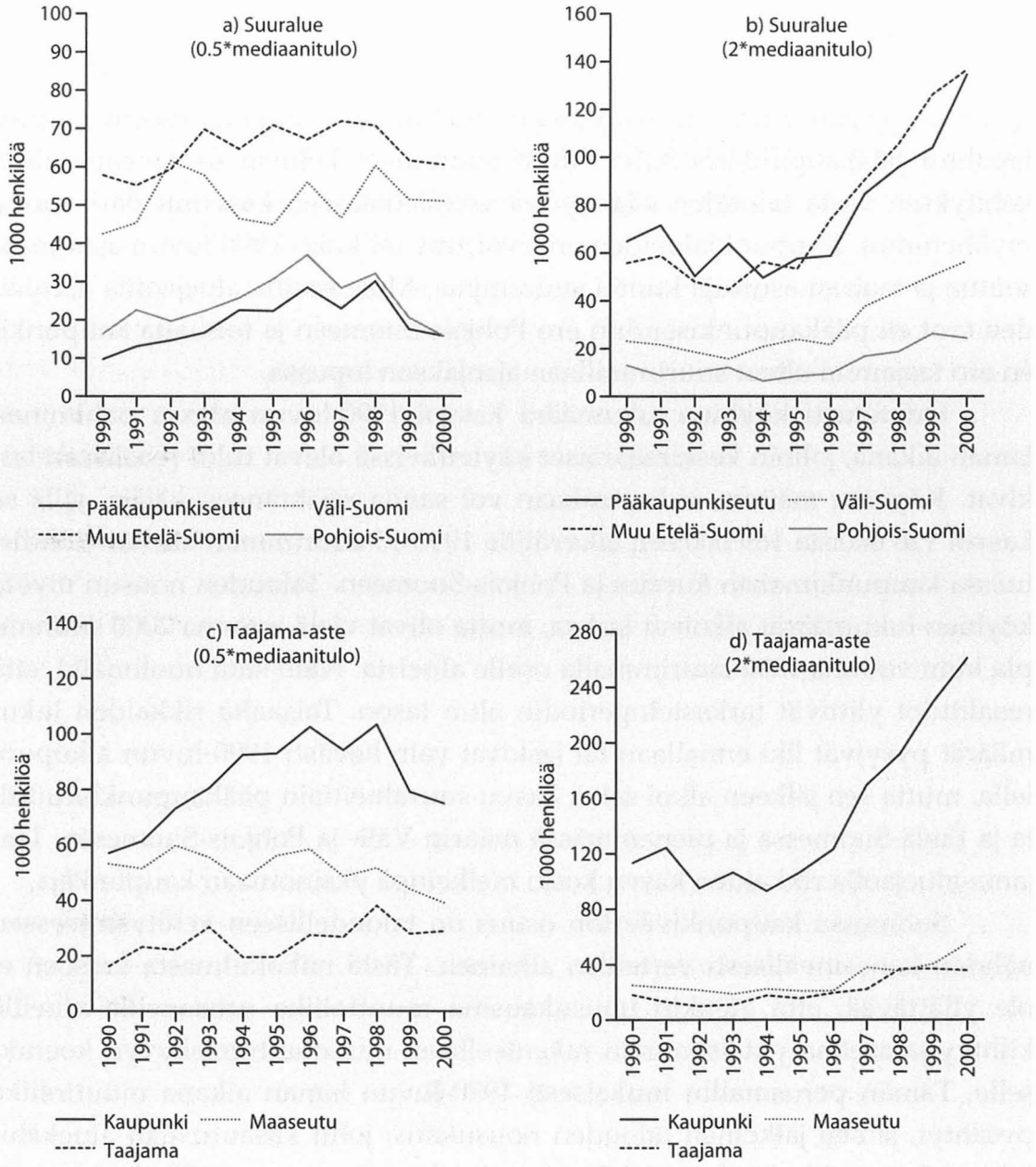
Rikkaiden lukumäärät pysyivät liki ennallaan tai laskivat hieman 1990-luvun vuosikymmenen alkupuolella. Sen jälkeen alkoi selvä kasvu Pääkaupunkiseudulla ja Etelä-Suomessa ja pienemmässä määrin Väli- ja Pohjois-Suomessa. Taajama-aluejaolla rikkaiden lukumäärän kasvu koski melkeinpä yksinomaan kaupunkeja.

LOPUKSI

Tässä kirjoituksessa tarkasteltiin käytettävissä olevien tulojen avulla alueiden välisiä suhteellisia tuloeroja ja alueiden sisäisiä tuloeroja. Tuloja ei deflatoitu alueellisilla hintaindekseillä, joita ei ollut saatavilla. Tällöin erityisesti pääkaupunkiseudun korkeat asumiskustannukset verrattuna muuhun maahan johtavat pikemminkin alueiden välisten erojen yli- kuin aliarviointiin. Aluejakoina käytettiin sekä pohjois-eteläsuunnassa maan jakavaa luokitusta että kaupungit, taajamat ja maaseutu -kolmijakoa. Tarkoituksena oli kartoittaa, johtiko 1990-luvun lama-aika tai sen jälkeinen nousukausi alue-erojen kasvuun tai alueiden sisäisten tuloerojen kasvuun. Tulonjakotarkastelua täydennettiin köyhien ja rikkaiden lukumäärien ja väestöosuuksien tarkastelulla.

Alueelliset tuloerot pysyivät yllättävän samanlaisina sekä laman aikana että sen jälkeen kummallakin aluejaolla tarkasteltuna. Tuloksia on vaikea, ellei mahdotonta tulkita niin, että alue-erot repesivät lamassa sen enempää pohjois-eteläsuunnassa kuin maaseudun, taajamien ja kaupunkien välillä. Vasta vuosituhannen viimeisinä vuosina, ei laman vaan korkeasuhdanteen myötä, alue-erot kasvoivat niin, että etenkin pääkaupunkiseudun ja Pohjois-Suomen ero revähti selvästi vuoden 1990 asetelmaa suuremmaksi. Kaupun-

Kuvio 15.3. Köyhien (alle puolet mediaanitulosta) ja rikkaiden (yli kaksi kertaa mediaanitulo) lukumäärät suuralue- ja taajama-astejaolla vuosina 1990-2000



Lähteenä on Tilastokeskuksen tulonjakotilaston palveluaineisto.

kien ja maaseudun suhteellisessa asemassa ei tapahtunut koko vuosikymmenellä juuri mitään. Sitä vastoin taajamaväestön suhteellinen asema, joka hieman parani 1990-luvun puolivälissä, on sen jälkeen heikentynyt ollen vuonna 2000 maaseutuväestön asemassa.

Alueellisia tuloeroja selvempi muutos 1990-luvulla näyttää tapahtuneen henkilöiden välisessä tulonjaossa sekä koko maan tasolla että alueiden sisällä, vaikka alku olikin yllättävä. Koko maassa ja eri aluejaoilla alueiden sisäisiä tulonjakoja tarkastellen 1990-luvun alkuvuosina ei ilmennyt juuri mitään muutosta: tulonjaon eriarvoisuutta mittaavien Gini-kerrointen arvot pysyivät ennallaan. Kun tuloerot vuonna 1993 alkoivat kasvaa, suurin muutos tapahtui pääkaupunkiseudulla. Muut suuralueet kokivat samansuuntaisen kehityksen vasta talouden käännyttyä voimakkaaseen kasvuun pari vuotta myöhemmin. Kaupunkialueiden eriarvoisuus oli koko 1990-luvun ajan maaseutua ja taajaan asuttuja kuntia suurempaa. Molemmilla aluejaoilla ääripäiden erot eli pääkaupunkiseudun ero Pohjois-Suomeen ja toisaalta kaupunkien ero taajamiin olivat suurimmillaan ajanjakson lopussa.

Odotetusti köyhien lukumäärä kasvoi 1990-luvun alussa pahimman laman aikana, jolloin keskimääräiset käytettävissä olevat tulot reaalisesti laskivat. Köyhien määrän ei kuitenkaan voi sanoa räjähtäneen käsiin, sillä se kasvoi 126 000:sta 169 000:een aikavälillä 1990-95 suurimman kasvun kohdistuessa kaupunkimaisin kuntiin ja Pohjois-Suomeen. Talouden nousun myötä köyhien lukumäärät alkoivat laskea, mutta olivat vielä vuonna 2000 suurempia kuin vuonna 1990 suurimmalla osalle alueista. Näin siitä huolimatta, että reaalitulot ylittivät tarkasteluperiodin alun tason. Toisaalta rikkaiden lukumäärät pysyivät liki ennallaan tai laskivat vain lievästi 1990-luvun alkupuolella, mutta sen jälkeen alkoi selvä kasvu suuralueittain pääkaupunkiseudulla ja Etelä-Suomessa ja pienemmässä määrin Väli- ja Pohjois-Suomessa. Taajama-aluejaolla rikkaiden kasvu koski melkeinpä yksinomaan kaupunkeja.

Suomessa kaupunkiväestön osuus on taloudelliseen kehitysasteeseen nähden kansainvälisesti vertaillen alhainen. Tästä näkökulmasta katsoen ei ole yllättävää, että etenkin nousukausina muuttoliike urbaaneille alueille kiihtyy ja asettaa yhteiskunnan rakenteellisen muuntautumiskyvyn koetukselle. Tämän perusmallin mukaisesti 1990-luvun laman aikana muuttoliike pysähtyi, ja sen jälkeinen talouden nousukausi johti kasautuvaan aluekehityksen. Lama ei kuitenkaan johtanut ensivaiheessa sen enempää alueiden välisten kuin niiden sisäisten tuloerojen kasvuun. Vasta talouden käännyttyä nousuun tuloerot koko maassa ja alueiden sisällä alkoivat kasvaa. Etenkin rikkaiden määrässä on tapahtunut voimakasta kasvua. Vasta vosituhannen loppuvuosina on havaittavissa alueiden välisten suhteellisten tuloerojen kasvua. Suurimmat tuloerot siis syntyivät, eivät niinkään laman aikana kuin talouden nousussa, ja toisaalta erot olivat merkittävämpiä ihmisten kuin alueiden välillä.

KIRJALLISUUTTA

Kalela, J., Kiander, J., Kivikuru, M-L., Loikkanen, H.A., Simpura, J. (2001): "Introduction" in Down from the Heavens, Up from the Ashes (eds. Kalela et al.), VATT Publications 27:6, Government Institute for Economic Research, Helsinki.

Kangasharju, A., Laakso, S., Loikkanen, H.A., Riihelä, M., Sullström, R. (2001): "Economic Crisis of the 1990s: What Happened to Regional Convergence and Inequality, and Housing Market Phenomena in Boom and Bust?" in Down from the Heavens, Up from the Ashes (eds: Kalela et al.), VATT Publications 27:6, Government Institute for Economic Research. Helsinki.

Kiander, J. ja Vartia, P. (1998): Suuri lama. Suomen 1990-luvun kriisi ja talouspoliittinen keskustelu. ETLA B 143. Taloustieto Oy. Helsinki.

Laakso, S. ja Loikkanen, H.A.: Kaupunkitalous. Kustantajalle toimitettu kirjan käsikirjoitus.

Lerman, R. ja Yitzhaki, S. (1984): "A Note on the Calculation and Interpretation of the Gini Index." Economics Letters 15, 363-8.

Loikkanen, H.A., Pekkarinen, J., Vartia, P. (toim.) (2002): Kansantaloutemme - rakenteet ja muutos, 3. uud. painos, Taloustieto Oy, Yliopistopaino, Helsinki.

Loikkanen, H.A., Rantala, A., Sullström, R. (1998): "Regional Income Differences in Finland during 1966-96", Government Institute for Economic Research, Discussion Papers 181, Helsinki.

Loikkanen, H.A., Riihelä, M., Sullström, R. (2000): "Tulo- ja kulutuserot kaupunkien, taajamien ja maaseudun välillä ja niiden sisällä" teoksessa Suomalaisen sosiaalipolitiikan alueellinen rakenne toim. Loikkanen, H.A., Saari, J., Sosiaali- ja terveysturvan keskusliitto, 16-40. Helsinki.

Riihelä, M., Sullström, R., Tuomala, M. (2001): On Economic Poverty in Finland in the 1990s. VATT Discussion Papers 264, Government Institute for Economic Research, Helsinki.

Riihelä, M., Sullström, R. (2001): Tuloerot ja eriarvoisuus suuralueilla pitkällä aikavälillä 1971-1998 ja erityisesti 1990-luvulla. VATT Tutkimuksia, 80.

Sullström, R. (1989): "Characteristics of the Lowest Income Quintile in Finland in 1981" in Essays on Income Distribution, Economic Welfare and Personal Taxation (eds. Robert Hagfors and Pentti Vartia). The Research Institute of the Finnish Economy, Series A 13, 129-163, Helsinki.

Tilastokeskus (1998): Alueluokitukset, Kunnat 1998, Käsikirjoja 28, Tilastokeskus.

Vartia, P. ja Vartia, Y. (1981): "Description of the Income Distribution by the Scaled F Distribution Model" in Statics and Dynamics of Income (eds. Klevmarken, N.A., Lybeck, J.A.), Tieto Ltd, Bristol.

LOPPUVIITTEET

[1] Kaupunkikehitystä ja kaupunkitaloustiedettä käsittelee koko laajuudessaan Laakson ja Loikkasen Kaupunkitalous -kirja, joka on tätä kirjoitettaessa kustantajalla.

[2] Suomen talouden viimeaikaisesta kehityksestä antaa peruskuvan Loikkanen, Pekkarinen ja Vartia (2002). Erityisesti 1990-luvun alun lamaa, sen syitä ja seurauksia on käsitelty laajasti mm. teoksissa Kiander ja Vartia (1998) sekä Kalela et al. (2001).

[3] Alueiden välisten ja sisäisten tuloerojen pidemmän aikavälin kehitystä sekä kulutustutkimus- että tulonjakoaineistojen avulla on käsitelty mm. tutkimuksissa Loikkanen et al. (2000) sekä Kangasharju et al. (2001) lamaan liittyvien kolmen aluetutkimuksen yhteenvetoartikkelissa (luku 13.3). Edellisessä sovelletaan taajama-aluejakoa vuosina 1966-1997 ja jälkimmäisessä suuraluejakoa vuosina 1971-1998.

[4] Pääkaupunkiseutu kattaa seuraavat kunnat: Espoo, Helsinki, Kauniainen ja Vantaa. Muu Etelä-Suomi koostuu seuraavista osista: Uudenmaan lääni (pl. pääkaupunkiseutu), Ahvenanmaan maakunta sekä vanhan läänijaon mukaiset Turun ja Porin lääni, Hämeen lääni ja Kymen lääni. Väli-Suomi kattaa vanhalla läänijaolla Mikkelin, Pohjois-Karjalan, Kuopion, Keski-Suomen ja Vaasan läänit. Vastaavasti Pohjois-Suomi koostuu Oulun ja Lapin lääneistä. Jaottelussa kunnat on vakioitu vuoden 1997 tasolle.

[5] Kaupunkimaisia kuntia ovat kunnat, joiden väestöstä vähintään 90 prosenttia asuu taajamissa tai suurimman taajaman väkiluku on vähintään 15 000. Taajaan asuttuja kuntia ovat kunnat, joiden väestöstä vähintään 60 prosenttia, mutta alle 90 prosenttia, asuu taajamissa, ja suurimman taajaman väkiluku on vähintään 4 000 mutta alle 15 000. Maaseutumaisia kuntia ovat kunnat, joiden väestöstä alle 60 prosenttia asuu taajamissa ja suurimman taajaman väkiluku on alle 15 000 sekä kunnat, joiden väestöstä vähintään 60 prosenttia, mutta alle 90 prosenttia, asuu taajamissa ja suurimman taajaman väkiluku on alle 4 000. Tilastollisessa kuntaryhmityksessä maantieteellisiä alueita ei vakioida vuosittain, vaan se kuvastaa väestön kasaantumisessa tapahtuneita muutoksia.

[6] Verotuksen ja tulonsiirtojen vaikutusta tulonjakaumaan on käsitelty mm. Riihelän ja Sullströmin (2001) tutkimuksessa, ja aihe on esillä myös aiemmissa alueellisia tuloeroja käsittelevissä tutkimuksissa (esim. Loikkanen, Rantala ja Sullström 1998), joissa sovelletaan rinnakkain useita tulokäsitteitä (tuotannontekijätulo, bruttotulo ja käytettävissä oleva tulo).

[7] Tutkijoiden keskuudessa logaritminormaali-, Pareto- ja F-jakauma ovat saaneet runsaasti huomiota poikkeuksellisten havaintojen ja siinä erityisesti tulojakauman ääripäiden tarkastelussa (vrt. Vartia ja Vartia, 1981).

16. Devaluation Cycles in Sweden and Finland

Ulf Jakobsson*

In both Finland and Sweden, the economic trend has been characterized by recurrent devaluations for a long period of time. Wages increasing too quickly, and cost crises with ensuing defensive or offensive devaluations have been well-known features of the economic-political picture on both sides of the Gulf of Bothnia. This has been such a regular pattern in both countries that it is possible to talk about a devaluation cycle. While the Finnish devaluation cycles started directly after WWII the pattern did not become regular in Sweden until the Bretton-Woods system broke up in the early 1970's. This is also the starting point for the period covered in this article.

However, the pattern of devaluation cycles came to an end due to the large, and in many ways similar, crises in both countries at the beginning of the 1990's. In connection with the crises, both countries changed exchange rate regimes; both went from an exchange rate target to a floating exchange rate with an inflation target.

In Finland, this regime was only a short-term intermediate solution before the country joined the ERM, and then the Euro. Sweden, on the other hand, has maintained the regime with a floating exchange rate and an inflation target for more than ten years. If there is a yes in the referendum in the fall (September 2003), Sweden will go the same way as Finland and adopt the Euro.

An important question for both countries is whether the forces creating the devaluation cycle under the previous regime are still active or whether the new regimes, possibly combined with a reform of the economies in the wake of the major crisis, mean that the requirements for continuous devaluations constitute a bygone stage for both countries.

If the answer to the question is no, that is if the underlying forces for devaluations still exist, there will be considerable problems in both regimes. Starting with the Swedish regime, the floating exchange rate can, in principle,

* Ulf Jakobsson is Director at the Research Institute of Industrial Economics (IUI) in Sweden.

currently satisfy a need for recurrent devaluations. In a very short perspective, the currency can depreciate, without violating the inflation target. However, relatively soon, recurrent depreciations will contribute to increasing the domestic rate of inflation. In order to defend the inflation target, the Riksbank will have to increase the interest rate so that inflation is kept down. In this scenario, the price for keeping down the inflation might be economic stagnation and growing unemployment. If, on the other hand, the Riksbank were to choose to accommodate the inflationary pressure and temporary or permanently abandoning the inflation target, it would risk losing its credibility. The price for this might also be very high, in the form of long periods with a considerable interest gap towards other countries until the confidence has been regained.

For a country that is part of the Euro area, the pressure to devaluate can naturally not be dealt with by new depreciations. In this regime, we might thus get too high a cost level, which must be regulated by a downward-adjustment of the relative wages. This might be a time-consuming process and must very possibly take place within the framework of economic stagnation and growing unemployment.

The question of whether the forces creating the recurrent devaluations in Finland and Sweden have now disappeared is thus of great importance for both countries. It is not, in any evident way, easier to deal with a recurrent pressure to devaluate within the framework of a floating exchange rate policy with a fixed inflation target than within the framework of a common European currency.

While few would deny the importance of this question, many might say that, in practice, it has already been answered by the actual trend. After the crisis in the early nineties, both Finland and Sweden have had a period of macroeconomic stability, which clearly differs from the turbulence in the period of devaluations, when there was a whole succession of macroeconomic crises.

Both countries have had considerable success in creating macroeconomic stability. However, it is also the case that this is largely due to the considerable depreciations in the wake of the crisis in the 1990's. In this context, Figure 16.1 is worth considering. Using a simple trend analysis as the starting point, it does not seem possible to claim that there has been a considerable break in the effective exchange rate trend for these two countries. After a strong downward deviation from the trend at the beginning of the 1990's, the countries are now once more approaching the trend. In this simple analytical

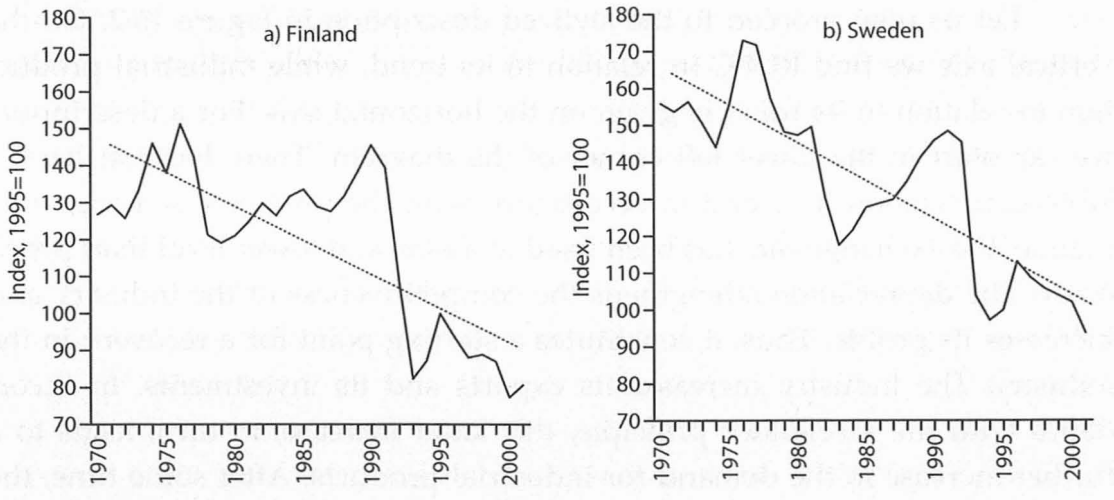
framework, the first signs of whether we are dealing with an actual break in the trend will only emerge in the next few years.

In this article, we will analyze this issue by a closer study of the Swedish and Finnish devaluation cycles. We shall discuss what have been the main driving forces of the cycles. Moreover, we will study the economic and political dynamics in the devaluation cycles. This will provide us with a background and an analytical framework for answering the initial question. We will return to this question at the end of the article.

There is a vast literature on devaluation cycles. Most of the research has been done with reference to one or several Latin American countries. Marion (1997) is one example of this. An early article on the devaluation cycle, which refers to the Finnish experiences after WWII is Korkman (1975). A description of the Swedish devaluation cycle is found in Jakobsson (1997).

The article is organised in the following way. In the next section we give a stylized description of the devaluation cycle. In the third section we analyze the actual experience of devaluation cycles in Sweden and Finland. In both countries the actual performance fits relatively well into the pattern outlined in the stylized description. In the fourth section we discuss the question of which are the driving forces behind devaluation cycle. On a general level we find a close interconnection between the political sphere and the economic sphere. On the one hand the stance of economic policy is an important factor driving the devaluation cycle. On the other hand the dynamics of

Figure 16.1. Relative Unit Labor Costs in Finland and Sweden (1970-2001)



Note: The source is OECD Economic Outlook.

the devaluation cycle gives important repercussions to the political sector. This is discussed in section five. In the last two sections we analyze the question of what brought the devaluation cycle to an end. We also come back to the question whether the forces that created the devaluation cycle are now laid to rest.

A STYLIZED DESCRIPTION OF THE DEVALUATION CYCLE

It might be suitable to start the analysis by a stylized description of the basic pattern of the devaluation cycle. At a very general level, the devaluation cycle is a way of describing a recurrent pattern in the interplay between the exchange rate trend on the one hand and the real economy on the other. The variables on which we will concentrate here are the relative unit labor cost (RULC) and industrial production, respectively. In both cases, we study their deviation from a trend. RULC is a measure of Sweden's real exchange rate, which determines the cost competitiveness of the industry, which is the reason why it seems natural to choose industrial production as the real variable.

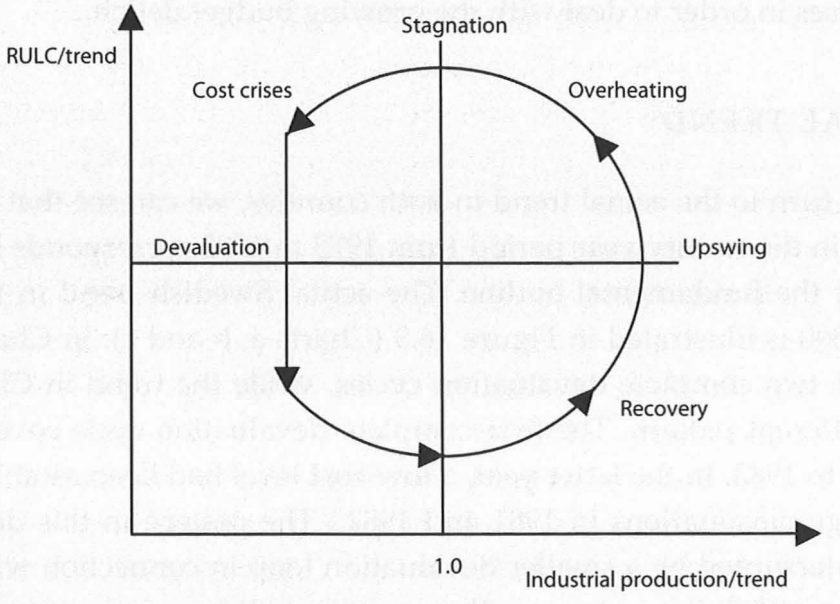
The cost of each unit produced is the ratio between wage costs and labor productivity. The relative wage cost per unit produced appears as the ratio between the ULC of the own country and a weighted average of the ULC of competing countries. With fixed exchange rates, the RULC is thus determined by the wage cost and the productivity in the country in relation to the corresponding variables in the other industrial countries. When the exchange rate for the currency of a country increases, so does its RULC, and the other way around when the exchange rate falls.

Let us now proceed to the stylized description in Figure 16.2. On the vertical axis we find RULC in relation to its trend, while industrial production in relation to its trend is given on the horizontal axis. For a description, we can start in the lower left corner of the diagram. There has just been a large depreciation here and in accordance with the previous exchange rate regime, the exchange rate has been fixed at a new and lower level than previously. The depreciation strengthens the competitiveness of the industry and increases its profits. Thus, it constitutes a starting point for a recovery in the industry. The industry increases its exports and its investments. In accordance with the accelerator principle, this latter increase, in turn, leads to a further increase in the demand for industrial products. After some time, the increase in industrial production leads to an increase in the demand for labor in the industry. In Sweden and Finland, this has usually led to wage inflation

taking off in the whole economy. Since the exchange rate is fixed, this leads to an increase in the relative wage cost in common currency. Increasing wages and growing employment also lead to an increase in domestic real income. Thus, there is also an increase in domestic demand in the economy. Another effect of this trend is that there is an increase in the tax base.

In the upper right-hand quadrant, we have, in summary, a situation where the economy starts to become overheated, with increasing wages, growing domestic demand and improved budget balances. At the same time, industry profits begin to fall, export volumes stagnate and industrial investments fall. Since the domestic part of the economy has now taken over the role as a driving force, the overheating can continue. The increase in private consumption is high. At the same time, the improved budget situation will lead to an increase in public consumption. The overheating, which is now basically nurtured by the domestic part of the economy, will finally lead to a cost crisis in the industry. The trigger might be a downturn in the international business cycle, which reinforces the problems in the exporting industry.

Figure 16.2. A Stylized Description of the Devaluation Cycle



The decline in the industry, with falling profits, decreasing employment and falling investments, affects the rest of the economy in various ways. First of all, it contributes to creating an increasing pessimism in the entire economy, which is reinforced if the decline coincides with a weaker international trend. Furthermore, the decline in the industry contributes to a worsened economic situation in the public sector, which contributes to a stricter, or at least more cautious, fiscal policy.

Decreasing employment in the industry, weakened domestic finances and growing budget problems serve to augment the crisis. This is further reinforced by the fact that traders in the currency markets get the impression of *déjà-vu*. The distrust leads to speculations against the currency and increasing interest rates, which adds further to the economic decline. Eventually a devaluation seems inevitable. Thus, in the terms of the figure, we have returned to the starting point.

In the two western quadrants, the budget situation of the public sector is strained by growing expenses due to a weaker labor market and falling incomes due to a weakened tax base. When we come to the south western quadrant, the loss in terms-of-trade after the depreciation becomes an additional factor, which further weakens the budget. In the south western quadrant, we can thus expect fiscal austerity packages with cuts in expenses and tax increases in order to deal with the growing budget deficit.

THE REAL TREND

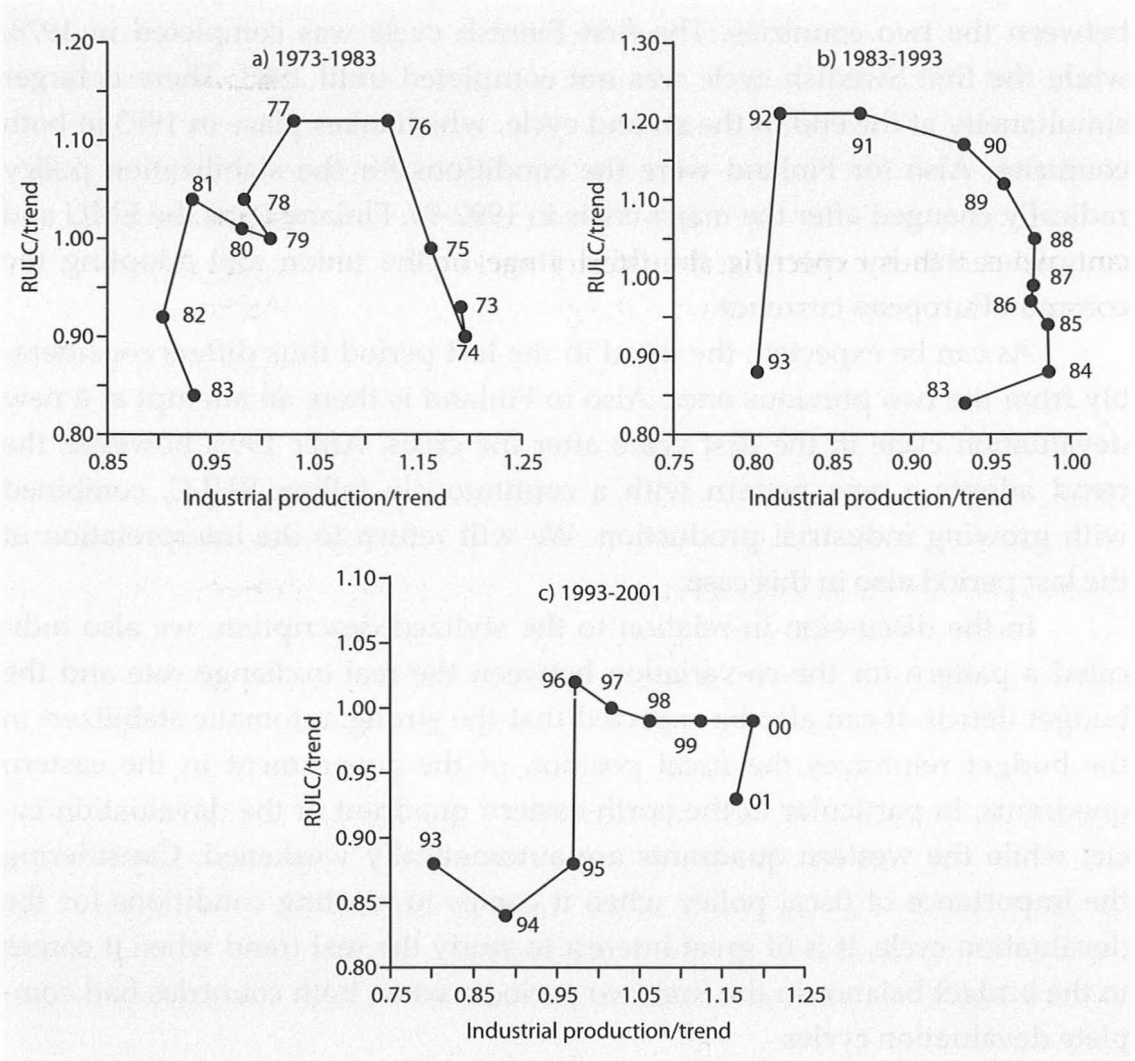
When we turn to the actual trend in both counties, we can see that the Swedish trend in the twenty-year period from 1973 to 1993 corresponds best to the pattern of the fundamental outline. The actual Swedish trend in the period 1973 to 2000 is illustrated in Figure 16.3 (Charts a, b and c). In Charts a) and b), we see two complete devaluation cycles, while the trend in Chart c) follows a different pattern. The first complete devaluation cycle covers the period 1973 to 1983. In the latter year, a low cost level had been established due to the large devaluations in 1981 and 1982.[1] The pattern in this devaluation cycle is interrupted by a smaller devaluation loop in connection with the devaluation in 1977. It was too small to create conditions for a recovery in the industry and accordingly, wage increases were modest in 1978. The strong fiscal expansion in 1979 and 1980 did, however, once more start driving the wage inflation. This paved the way for the devaluations in 1981 and 1982, amounting to 10 and 16 per cent, respectively. Despite the experience from the previous devaluation, the government was very optimistic that the large

devaluations in 1981 and 1982 would lead to the structural change that constituted the explicit goal of the devaluations.

Instead, as appears from Chart b), 1983 constituted the starting point of a devaluation cycle, which almost perfectly corresponds to the pattern in the fundamental outline in Figure 16.2. It came to an end with the exchange rate crisis and the large depreciation in 1992.

Then, the conditions for the stabilization policy were radically changed. First of all, Sweden abandoned the policy with a fixed exchange rate. The new exchange rate regime was a floating currency with an *inflationary target.* Furthermore, the Swedish Riksbank was restructured, so that its board became independent of the government. A ceiling for government

Figure 16.3. Devaluation Cycles in Sweden



Note: The source is OECD Economic Outlook.

spending was introduced in the fiscal policy. Moreover, a long-term goal was introduced for the trend of the budget balance. With the entry into the EU in 1996, Sweden also became a member of the EMU; the independence of the Riksbank and the budget goal can be considered as direct consequences of this.

The observed pattern in the period after 1993 also very obviously deviates from the both previous time periods. Initially, we do find an attempt at a new devaluation cycle. After 1996, the trend adopts another pattern, however. We shall return to the question of whether this means that the old pattern has now entirely disappeared.

The real trend in Finland is illustrated in Figure 16.4 (Charts a, b and c). Also in this case, there are obvious cycles in the first two time periods, even if the deviations from the pattern in the stylized description are somewhat larger than for Sweden. The division into periods also differs somewhat between the two countries. The first Finnish cycle was completed in 1978, while the first Swedish cycle was not completed until 1983. There is larger simultaneity at the end of the second cycle, which takes place in 1993 in both countries. Also for Finland were the conditions for the stabilization policy radically changed after the major crisis in 1992-93. Finland joins the EMU and completes this by entering the third stage of the union and adopting the common European currency.

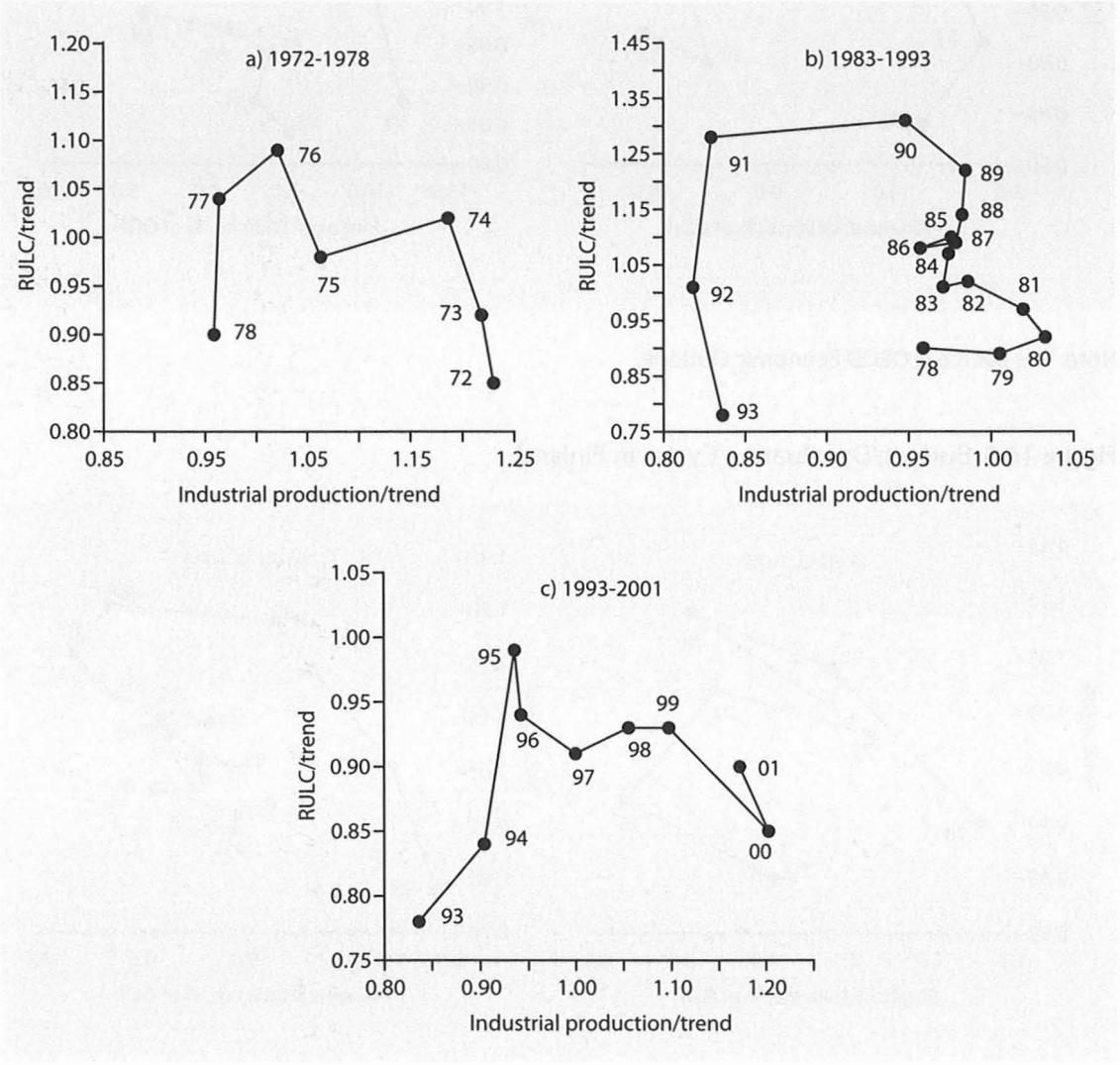
As can be expected, the trend in the last period thus differs considerably from the two previous ones. Also in Finland is there an attempt at a new devaluation cycle in the first years after the crisis. After 1995, however, the trend adopts a new pattern with a continuously falling RULC, combined with growing industrial production. We will return to the interpretation of the last period also in this case.

In the discussion in relation to the stylized description, we also indicated a pattern for the co-variation between the real exchange rate and the budget deficit. It can also be expected that the strong automatic stabilizer[2] in the budget reinforces the fiscal position of the government in the eastern quadrants, in particular in the north-eastern quadrant of the devaluation cycle, while the western quadrants are automatically weakened. Considering the importance of fiscal policy when it comes to creating conditions for the devaluation cycle, it is of great interest to study the real trend when it comes to the budget balance in the first two periods, when both countries had complete devaluation cycles.

The real trend for this variable in Sweden is shown in Figure 16.5 (Charts a and b) and for Finland in Figure 16.6 (Charts a and b). In the case of

Sweden, we can see that the pattern for both cycles shows the expected trend. Most likely, the correlation between the automatic stabilizer in the budget on the one hand and discretionary economic policy on the other hand has been a driving force in the devaluation cycle. When the government budget shows a growing surplus in the upturn, this contributes to increasing the general euphoria characterizing the latter part of the upturn in the devaluation cycle. The growing surplus is considered as a signal of there being scope to commit to new expenditures or, alternatively, tax cuts. Part of the explanation for the drastic deterioration in the budget balance that we see in the western quadrants might be that fiscal policy commitments made at the end of the upturn now have full impact.

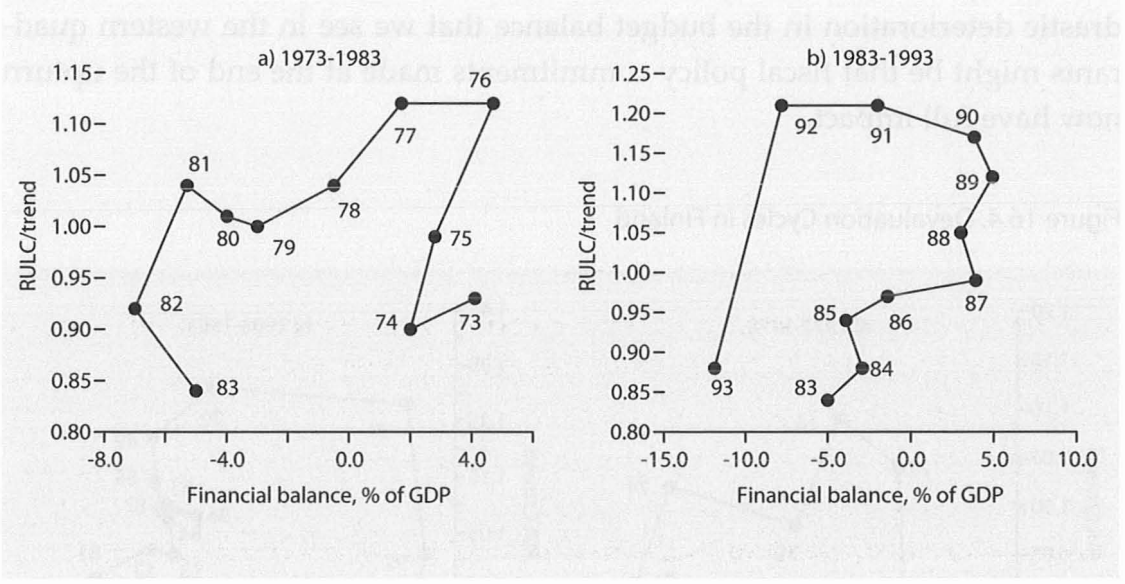
Figure 16.4. Devaluation Cycles in Finland



Note: The source is OECD Economic Outlook.

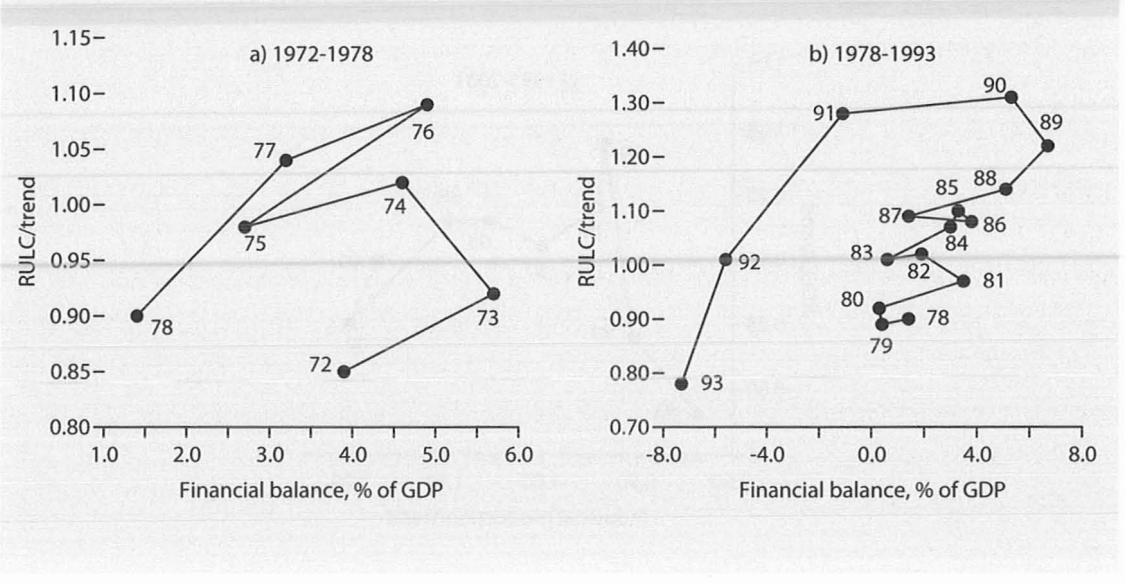
The Finnish budget cycles also largely follow the pattern that has been outlined. In the first cycle, 1975 constitutes a deviation, which is difficult to explain. 1976, which constitutes the peak year in the north eastern quadrant, also gives the largest budget surplus. Then, in accordance with the stylized description, there is a drastic weakening when the cycle enters a downturn. If

Figure 16.5. Budget/Devaluation Cycles in Sweden



Note: The source is OECD Economic Outlook.

Figure 16.6. Budget/Devaluation Cycles in Finland



Note: The source is OECD Economic Outlook.

we go to the second time period, the budget cycle almost exactly corresponds to the stylized description.

THE DRIVING FORCES OF THE DEVALUATION CYCLE

We shall now deal with the question of what have been the driving forces in creating the devaluation cycles in both countries. There will be no general and complete statistical analysis of this question. Instead, we will point at two underlying factors that have been common for both countries and argue for their having been of decisive importance for the devaluation cycle.

Unemployment Has, for a Long Period of Time, Constantly Been Pressed Down below NAIRU

An almost obvious observation in connection with the devaluation cycle is that the wage costs in both Sweden and Finland have increased faster in the period 1970 to 1992 than in the other industrial countries. Depending somewhat on how the comparison group is defined, Swedish wages in local currency have increased about 35 per cent faster than in the average competitor country. The corresponding figure for Finland is even larger. Since productivity has not grown so much faster than in other industrial countries, it is obvious that this trend has contributed to decreasing the exchange rate.

The next question is then why wages have increased faster in these two countries than in the other industrial countries. An obvious reason might seem to be that wage formation has been inefficient. This "obvious reason" does, however, conflict with a considerable amount of research in this area. In the 1980's, several researchers claimed that wage formation worked better in Sweden than in the other industrial countries. In a number of studies, it appeared that Sweden was one of the countries in the OECD area where wage formation worked best, in the sense of the Swedish labor market showing considerable flexibility in real wages.[3] In an OECD study[4] belonging to this literature, both Sweden and Finland come out very favourably in an international comparison. As we have seen earlier in this article, the adjustment in real wages however was not achieved by low wage increases but by the devaluations being even larger than the wage increases. It is not obvious to all critics that such a trend should be interpreted as a country having a well-functioning wage formation.

Most likely, deficient wage formation should not be given too large a part of the blame for the trend creating the devaluation cycle. Instead, it is probably the case that due to the economic policy, wage formation was subject to more pressure than it could handle. One way of seeing this is to compare the unemployment figures in both countries with the estimates that have been made for NAIRU in the respective country. In Sweden, we find estimates for NAIRU in the interval of 3 to 3.5 per cent in the period before 1992[5]. For 1993 Forslund (1995) reports values between 4 and 6 percent. In Finland, the values of NAIRU have also increased over time albeit on a somewhat higher level than in Sweden. According to the estimates of Alho (2002), NAIRU took values around 4 percent in the period 1970-1983. Then the values increased until they reached 8 percent in the early 1990's. After that NAIRU has remained at that level.

If we compare the unemployment figures in these countries to these estimates, we find for Sweden that in each single year in this period, unemployment has been below the estimated value for NAIRU. In Finland unemployment has been below or just at NAIRU during the whole period. This is a sign that a too expansionary economic policy in the two countries has been of crucial importance for creating the devaluation cycle.

The question is thus how both countries could stick to a policy that was so obviously in conflict with such central goals for economic policy as low inflation and a fixed exchange rate. Considering Sweden, the fact that the government was assuming the entire responsibility for maintaining full employment is certainly an important explanation for the implemented policy. Full employment was moreover interpreted as a level of unemployment considerably below NAIRU. While the government was assuming responsibility for full employment, the labor market parties were, in practice, responsible for the inflation. It was also the case that the concept of NAIRU was not accepted at the political level. In the 1970's and 1980's, there was a strong belief that the strength of the Swedish labor movement could, in principle, make it possible to combine unemployment figures below NAIRU with non-inflationary wage increases. This view was manifested in a large number of political agreements, where the government, in a number of political concessions to the trade unions, tried to make these commit to low wage demands which would, in turn, lead to low overall wage increases. In Sweden, the so-called "Haga[6]-rounds" in the 1970's constituted a very clear example of this structure of the policy. On a permanently overheated labor market, this was a task that could not be handled by the trade unions. Accordingly, Sweden was heading towards a trend where too large wage increases became a threat to

full employment. This threat, in turn, was averted by a more expansionary fiscal policy and later on, new devaluations. This, in turn, meant that the ground was paved for continued wage inflation, which once more had to be accommodated by an expansionary fiscal policy and new devaluations, and so on. Finland's history of centralized tripartite wage agreements actually is longer than Sweden's[7]. Also here the ambition was to "beat the NAIRU" by policy agreements with the unions.

A Deteriorating Trend in the Real Exchange Rate

The devaluations in Sweden and Finland have not only served as adjustments of an inflationary wage trend. They have also constituted an adjustment to a decreasing relative wage-level in both countries. Already in Figure 16.1, we saw that there has been a deteriorating trend in the real exchange rate ever since the beginning of the 1970's. This indicates that there has been a *continuous decrease in the underlying competitiveness* in both Sweden and Finland. In order to deal with competition in the existing economic structure, there has been a need for a continuous downward adjustment of the relative wages. This trend has been most apparent in Sweden. The situation there is illustrated by the fact that the depreciation that became unavoidable in Sweden in 1992, started from a real exchange rate level about 25 per cent below the level at the time of the devaluation in 1977. If considering the entire time period, the real exchange rate in Sweden has fallen by 35 per cent. If we assume the wage increases in Sweden in this entire period to have been at the same level as in the other industrial countries and that this, in turn, has not affected the underlying competitiveness, Sweden would still have had to devaluate by around 30-40 per cent in that period in order to achieve the same cost competitiveness as today. The corresponding figures for Finland are of same order of magnitude.

When analyzing the reasons for this trend, it is important to note that it cannot be explained by the wage costs having been adjusted to a weak productivity trend. The real exchange rate trend emerges as the difference between the wage formation trend and the productivity trend. Thus, the productivity trend has already been weighted into the measure. Instead, the trend must be interpreted as the competition for the countries' products having becoming stiffer throughout, which has led to a downward adjustment of prices and thus, also costs. We shall return to possible explanatory factors for this in a later section.

THE POLITICAL DYNAMICS OF THE DEVALUATION CYCLES

We found in the previous section that one of the main driving forces behind the devaluation cycle was that unemployment was constantly pressed below NAIRU. The devaluation cycle was part of an economic policy that generally was too expansionary. On a very general level the devaluation cycles seem to have their origin in the political sphere. This is an indication that there is a connection between the devaluation cycles and the electoral cycles.

In the 1970's, William Nordhaus (1975) wrote a very influential article introducing a theory of *electoral cycles.* The main argument of this article was that in its attempt to gain voters in a coming election, the present government might be expected to carry out an excessive expansionary policy in order to increase its possibilities of being re-elected. The excesses before the election will then need to be corrected by a tightening after the election. This theory has been developed in a number of different articles and it has been empirically tested, with varying results.

An alternative theory developed by Douglas Hibbs (1977) among others, is that the nature of a policy before an election is determined by the political color of the government. A left-wing government can be expected to pursue an expansionary policy before the elections in order to show good unemployment figures. According to this theory, a conservative government can instead be expected to pursue a restrictive policy in order to show low inflation figures. This theory has also constituted the starting point of several empirical studies.

The theories have generated studies both on Swedish and Finnish data. A recent study on Finland is Nyberg (2000). On Sweden, studies have been made by Hibbs & Madsen (1981) and Jonung (1985). A general result of the studies is that elections are important for the timing of stabilization policy.

One can however not conclude that devaluation cycles simply are another incarnation of electoral cycles. There are important differences between the electoral cycles and the devaluation cycles we are analyzing in Finland and Sweden. There is e.g. an obvious difference in periodicity. The electorate cycle spans the election period, which both in Sweden and Finland is four years. In the theory of the electorate cycle the policy gets expansionary just before the election and more restrictive after the election. In the exposition of the devaluation cycle we have seen that the policy stance have been too expansionary, during periods of several years.

We have also found that an important feature of the of the devaluation cycle is, that the dynamics of the cycle tends to set the rhythm of the expan-

sions and restrictions, respectively. Figure 16.5 and Figure 16.6 give a clear indication of this feature.

The electoral cycle is, in theory shaped entirely by the political sphere. In analyzing the devaluation cycle we find important feedbacks from the economic sphere to the political sphere. We have just given one example of that. The dynamics of the devaluation cycle is also an important part of the environment in which the political parties are acting. We shall illustrate this with a reference to the Swedish case.

Since the beginning of the 1970's, there have been two periods with right of center governments in Sweden: 1976-1982 and 1991-1994. Both these government periods have thus started at the peak of the devaluation cycle and ended near its lowest level. The social democratic governments have thus administered the phases of recovery until about a year before the start of the cost crisis. For the fiscal policy, this means that the years that the right of center parties have been in power, have been years with large and growing budget deficits while the years when the Social Democrats have been in power, have been years with a good budget situation. In particular at the end of the recovery, the budget has shown a large surplus.

However, this cannot simply be interpreted as the Social Democrats generally carrying out a stricter fiscal policy than the right of center parties. Instead, as we have pointed out earlier, the policy has, in particular at the end of the recovery, been very expansionary, which has been hidden by the strong automatic feature in the budget.

The right of center budget trend has, on the other hand, been characterized by the decline, since the automatic feature has been working in the other direction. At the same time, the policy in the years when the right of center parties were in power must also be considered to be expansionary. In the first right of center government period, the policy from 1979 and onwards was very expansionary. Also in the second right of center government period was the decline met by very strong discretionary increases in public expenditure. It is, however, difficult to find any support for the theory that left of center governments would carry out a more expansionary policy than right of center governments on basis of the Swedish experiences. Using the period 1970 to 1992 as a starting point, it should also be difficult to argue that the opposite would hold. In the period of consolidation after 1992, a balancing of the budget, in combination with stricter regulations for fiscal policy, was implemented by a social democratic government.

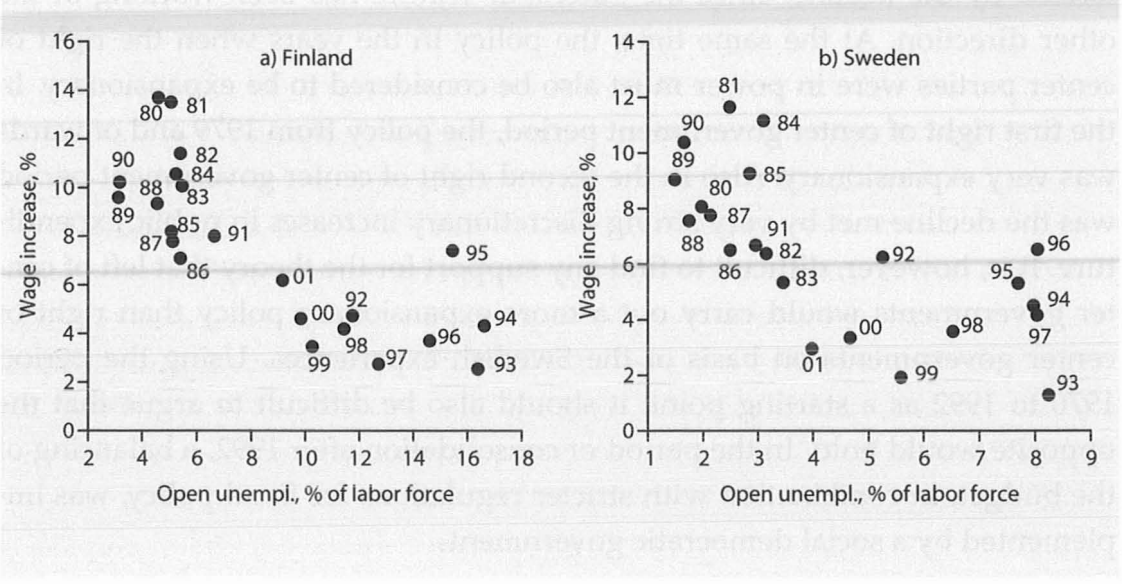
THE END OF THE DEVALUATION CYCLES

In both Finland and Sweden, the devaluation cycles came to an end at the beginning of the 1990's. In both countries, a very deep macroeconomic crisis constituted the end of this era. GDP fell considerably in both countries and the unemployment figures rose very quickly to levels that would have been considered as entirely impossible, from a political point of view, only about a year before the crisis. The unemployment trend appears from Figure 16.7 (Charts a and b).

It is obvious that there were more profound reasons for the crisis than the cost crises that both economies had been subject to on a number of previous occasions. For both countries, there are a number of analyses on the reasons for the crisis[8]. In Finland, not least Pentti Vartia's work[9] has contributed to increasing the knowledge of the background to the deep crisis.

There is no space here to reproduce or even less to further develop these analyses. An interesting observation, however, is that in the 1980's, both countries underwent a considerable change in the direction of being liberal market economies. In particular, the capital markets were deregulated and interest rate regulation was abolished. This meant that the entire capital markets in both countries were allowed to, or were forced to, work according to the principles of a market economy. This had profound implications for a number of businesses building on the regulated capital market; in particular

Figure 16.7. Wage-Increase and Unemployment in Finland and Sweden (1980-2001)



Note: The sources are OECD Economic Outlook and US Department of Labor, Bureau of Labor Statistics.

the housing sector, the construction sector and the banking sector. These lines of business all had considerable problems in both countries during the crisis in the early 1990's. The requirements for profitability in investment activities also became stricter in the entire economy, when the real interest rate went from a value around minus five to plus five in both economies. In Finland, the additional decline in the trade with the eastern countries meant that large parts of the export sector had to adapt to market-economic conditions.

A parallel with the *transition crises* that many eastern economies underwent at that time is close at hand. The trend of the economic activities in both countries also followed the pattern of a deep decline and an ensuing recovery, which was typical of the beginning of the transition phase of the East European economies.

The extent of the decline was most probably also due to the costs for the defense of the currency before the fixed exchange rate finally had to be abandoned. This cost was most certainly highest in Sweden, where the exchange rate was defended for a longer period of time and at considerably higher interest levels than in Finland.

In our context, the depth and the extent of the crisis do raise the question of to what extent too high a cost level did contribute to the crisis. Would the crisis have evolved in any significantly different way, if the adjustment to a lower exchange rate had come at a considerably earlier stage? My own view is that the difference would not have been of great importance for Finland. For Sweden, as we have just seen, the costs of defending the krona in the form of very high interest rates for an extended time contributed to a deepening of the crisis.

In retrospect, it is obvious that the defense of the Swedish krona was futile after Britain had left the ECU. If Sweden had devaluated at this point in time, the long period of defending the krona with exorbitant interests would not have occurred and the crisis in the financial sector would probably not have been as deep as it turned out to be either. This, in turn, would have paved the ground for a more modest decline in the entire Swedish economy.

A very important result of the deep decline for both Finland and Sweden was that both countries experienced a swift and strong disinflation. Only in one year during the 1980's did the Finish wage increases fall below 8 percent. As late as in 1989, the wage increases amounted to about 10 percent. Two years later, in 1991, they amounted to about 2 percent. The picture is similar in Sweden, even if the decline in the wage increases is less apparent here. The decline in the labor market was not as deep in Sweden as in Finland either. The question to be asked is whether a less deep decline would

have created a sufficiently strong disinflation to break the devaluation cycle in the way it actually occurred.

STABILITY FOREVER?

The devaluation cycle was connected to what in retrospect appears to be a failed fixed exchange rate regime. Both Finland and Sweden abandoned this regime in connection with the crisis in the 1990's. Thus, the cyclical pattern we have surveyed here will hardly recur. However, it might naturally be the case that those forces that created the devaluation cycle still exist or will reappear. Finland and Sweden would thus once more risk macroeconomic instability. However, it would most likely manifest itself in ways different from the devaluation cycles studied in this article. Finally, we will discuss this question by returning to the factors that we have earlier pointed out as being the driving forces of the devaluation cycle.

The first question is then whether the *risk of wage inflation* has now been averted. In connection with Figure 16.7, we found that the rate of the wage increases in both countries was considerably dampened after the crisis in the 1990's. In both countries, unemployment had also been driven up above the NAIRU. This has been most obvious in Finland. Thus, there are also conditions for a more stable wage formation process than earlier. The important question is naturally whether this situation will last. It can be found that both countries have made changes in the regulations and institutions, which counteract new outbreaks of wage inflation. For Finland, the adoption of the Euro means that all ideas, that high wage increases will be *bailed out* by devaluation, have disappeared. This should have a cooling effect on both wage formation in the labor market and a government that might have planned to reduce unemployment below NAIRU by an expansionary policy. In addition, the stabilization pact within the Euro should counteract this kind of excessive fiscal policy, which earlier contributed to driving up the price- and wage inflation.

Within the framework of the second stage of the EMU, Sweden has also carried out a number of changes that increase the macroeconomic stability. The most important ones are that the Riksbank has obtained an independent position according to the constitution, and that new regulations for the fiscal policy have been introduced. Thus, by changing the regulations for parliamentary decisions, the government position has been strengthened when it comes to controlling government expenditures. A long-term ceiling for government expenditures has been introduced as part of the same reform.

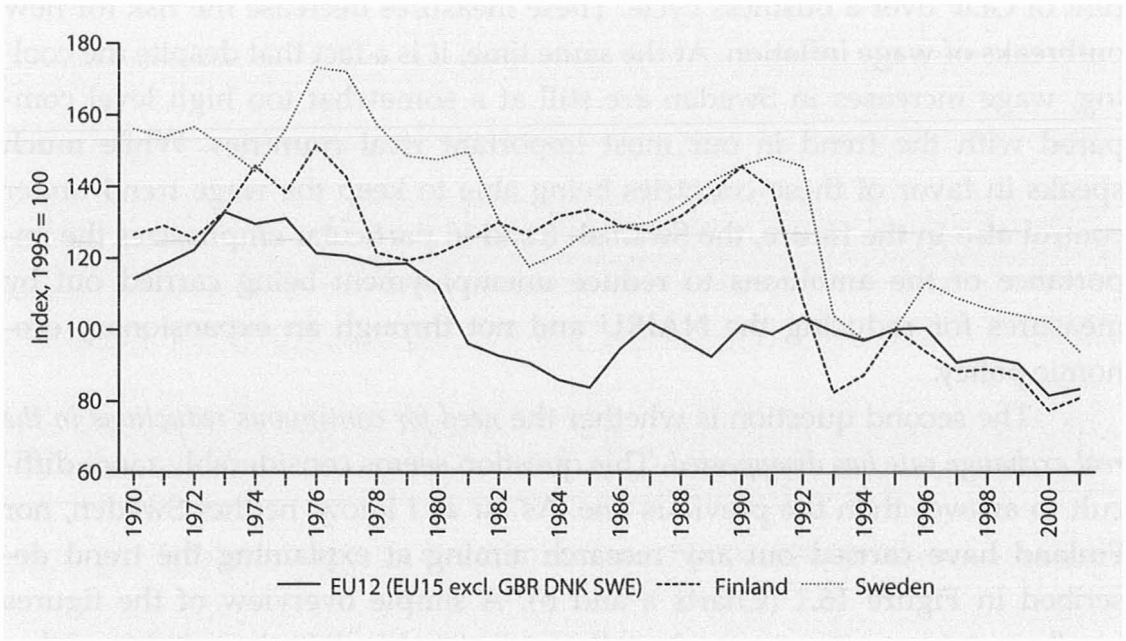
Moreover, there is a target for the budget balance to be on average two percent of GDP over a business cycle. These measures decrease the risk for new outbreaks of wage inflation. At the same time, it is a fact that despite the cooling, wage increases in Sweden are still at a somewhat too high level compared with the trend in our most important rival countries. While much speaks in favor of these countries being able to keep the wage trend under control also in the future, the Swedish trend in particular emphasizes the importance of the ambitions to reduce unemployment being carried out by measures for reducing the NAIRU and not through an expansionary economic policy.

The second question is whether the *need for continuous reductions in the real exchange rate has disappeared*. This question seems considerably more difficult to answer than the previous one. As far as I know, neither Sweden, nor Finland have carried out any research aiming at explaining the trend described in Figure 16.1 (Charts a and b). A simple overview of the figures hardly provides any material for talking about a break in the trend for either of the countries. In both countries, the yearly decrease in the real exchange rate is about 1.5 per cent per year. In a historical perspective, this trend is due to the recurrent devaluations/depreciations we have analyzed in this article.

For Finland, the possibility of making independent depreciations no longer exists. The question is thus whether the trend indicates that adjusting the Finnish economy to the adoption of the Euro will be a great problem. A study of Figure 16.8 on page 306 does not indicate that this would necessarily be a great problem. Considering the Euro countries as a whole, the real exchange rate has also shown a falling trend. Seen over the entire period, the decline amounts to 1.2 per cent per year. The difference between 1.5 and 1.2 percent could hardly be a cause for concern. In adopting the Euro, the Finnish economy can be expected to become increasingly integrated in the Euro area. Thus, a further convergence in the real exchange rate trend can also be expected.

The analysis for Sweden is similar. The difference in the trend of the real exchange rate is small. Thus, it should be possible for the krona to keep in step with the Euro in the long run if Sweden can keep wage inflation under control. If Sweden adopts the Euro, the Swedish situation in this respect will be similar to that of Finland. Also in the case of Sweden can we in this case expect a convergence of the real exchange rate to the Euro area.

Figure 16.8. Relative Unit Labor Cost in Finland, Sweden and the Euro-Area (1970-2001)



Note: The source is OECD Economic Outlook.

REFERENCES

Alho, Kari E.O. (2002): "The Equilibrium Rate of Unemployment and Policies to Lower It: The Case of Finland", ETLA Discussion Paper No. 839.

Braconier, H. and Holden, S. (1999): "The Public Budget Balance, Fiscal Indicators and Cyclical Sensitivity in the Nordic Countries", TemaNord, 1999:575, Nordic Council of Ministers, Copenhagen.

Eriksson, Tor, Suvanto, Antti and Vartia, Pentti (1990): "Wage Formation in Finland", in *Wage Formation and Macroeconomic Policy in the Nordic Countries*, Lars Calmfors (ed.), SNS Förlag and Oxford University Press, pp. 189-235.

Forslund, A. (1995): "Unemployment: Is Sweden Still Different?", *Swedish Economic Policy Review*, Vol. 2, pp. 15-58.

Hibbs, D.A. Jr. (1997): "Political Parties and Macroeconomic Policy", *American Political Science Review*, Vol. 71, pp. 1467-1487.

Hibbs, D.A. Jr. and Madsen, H.J. (1981): "The Impact of Economic Performance on Electorate Support in Sweden, 1967-1978", *Scandinavian Political Studies*, Vol. 4, No. 1, pp. 33-50.

Holmlund, B. (1990): "Svensk lönebildning – teori, empiri, politik". Bilaga 24 till Långtidsutredningen 1990.

Jakobsson, Ulf (1997): "Den svenska devalveringscykeln", *Ekonomisk Debatt*, Vol. 25, No. 3, pp. 141-154.

Jonung, Lars (1985): "Business Cycles and Political Changes in Sweden", *Skandinaviska Enskilda Banken Quarterly Review*, 2, pp. 26-39.

Kiander, Jaakko and Vartia, Pentti (1996): "The Great Depression of the 1990s in Finland", *Finnish Economic Papers*, Vol. 9, pp. 72-88.

Korkman, Sixten (1978): "The Devaluation Cycle", *Oxford Economic Paper*, Vol. 30, No. 3, pp 357-376.

Lindbeck, Assar (1997): *The Swedish Experiment*, SNS Förlag.

Marion, Nancy P. (1997): "Devaluation Cycles and Adjustment Costs", in *International Trade and Finance: New Frontiers for Research*. Essays in Honor of Peter B. Kenen, Cambridge University Press, pp. 339-360.

Nordhaus, W.D. (1975): "The Political Business Cycle", *Review of Economic Studies*, Vol. 42, pp. 169-190.

Nyberg, Martti (2000): "Interaction between The Economy and Politics in Finland. A Study on Popularity of Parties and Political Business Cycles", Helsinki School of Economics and Business Administration, Acta Universitatis Oeconomicae Helsingiensis, A-170.

OECD (1989): Economies in Transition. Structural Adjustment in OECD Countries. OECD, Paris.

Wadensjö, Eskil (1987): "Labour Market Policy and Employment Growth in Sweden", *Labour*, Vol. 1 No. 3.

ENDNOTES

[1] Both devaluations took place in the fall, and thus only had a full impact on the figures for the whole year in the following year.

[2] For an estimation of the automatic stabilizers in the Nordic countries see Braconier & Holden (1999).

[3] See, e.g. Holmlund (1990).

[4] OECD (1989).

[5] See Holmlund (1990) and Wadensjö (1987).

[6] The negotiations between the government and the social partners were carried out at the government castle of Haga.

[7] For a review of the wage agreements in Finland after WWII see Eriksson, Suvanto and Vartia (1990).

[8] Such an analysis is included in Lindbeck (1997).

[9] See e.g. Kiander and Vartia (1996).

17. Talouspoliittinen keskustelu ja suuri lama

Jukka Pekkarinen*

Eräs pitkäaikaisimpia talouspoliittisia keskustelukumppaneitani on varmasti ollut Pentti Vartia. Muistan hänet hyvin 1970-luvun alusta yliopiston kansantaloustieteen laitoksen tutkijatiloista Katajanokalta. Vartia oli pestattu diplomi-insinöörin ja valtiotieteilijän paperit taskussaan professori Pentti Pöyhösen integraatioryhmään. Siitä hän kohta siirtyi apulaisprofessoriksi kansantaloustieteen laitokselle ennen kuin aloitti pitkän työrupeamansa Elinkeinoelämän Tutkimuslaitoksessa.

1970-luvun alun Katajanokalla Pentti oli muistikuvieni mukaan hyvin samankaltainen kuin aina myöhemminkin. Hänen ulkonäkönsäkään ei ole juuri muuttunut. Nuorekas herrasmiehen olemus on edelleen sama. Mutta vielä tuoreempia ovat muistikuvani Pentistä keskustelevana ekonomistina. Myöhempinä yhteistyömme, keskustelujemme – ja väliin väittelyjemmekin – hetkinä nuo muistikuvat ovat vain vahvistuneet.

Pentti Vartia on aina ollut innokas keskustelija. En muista ajan puutteen estäneen häntä koskaan ainakin hetkeksi pysähtymästä kasvokkain käytyyn mielipiteiden vaihtoon jostakin ajankohtaisesta talouspoliittisesta tai talousteoreettisesta kysymyksestä. Keskustelijana Pentille on ollut ominaista hyväntahtoisuus ja myönteisyys. Hän löytää toisen osapuolen näkemyksistä aina jonkin hyvän puolen eikä koskaan tyrmää niitä oitis. Hän tuntuu aina tulevan askelen vastaan.

Myös hyvä pelkistäjä Pentti Vartia on. Pelkistämisessä voi nähdä myös hänen insinöörin ja tutkijan koulutuksensa. Kaikkein mieluiten Vartia kiteyttää oman näkemyksensä joksikin yksinkertaiseksi tilastolliseksi riippuvuudeksi tai analyyttiseksi kuvioksi. Näitä hän sitten muistaa toistaa eri yhteyksissä. 1970- ja 1980-lukujen vaihteen tienoilla Pentti Vartian takinliepeessä seurasi eri piireihin keskustelu siitä, tulisiko esimerkiksi liikevaihtoveron korotukseen liittyvän hintojen nousun vaikutus kuluttajien ostovoimaan suhteuttaa kulutusmenoihin vai käytettävissä oleviin tuloihin. Määritelmällisenä erona näiden kahden menettelyn välillä on säästämisen kohtelu. Määrittely-

* Jukka Pekkarinen on Palkansaajien tutkimuslaitoksen johtaja.

eron tosiasiallinen merkitys piilee siinä, että liikevaihtoveron korotuksen ostovoimavaikutukset ovat vähemmän regressiivisiä, jos ne suhteutetaan kotitalouksien kulutusmenoihin niiden käytettävissä olevien tulojen sijasta. Tämä johtuu yksinkertaisesti siitä, että kuluttamatta jätetty osa käytettävissä olevista tuloista määritellään säästämiseksi ja sen osuus kotitalouksien tuloista kasvaa tulojen kasvun myötä.

Kollegojen on helppo palauttaa mieliin monia vastaavia esimerkkejä Pentti Vartian pedagogisista pelkistyksistä. Eräs niistä liittyi 1980-luvun keskustelussa devalvaatiosykliin: jos devalvaatio suoritetaan keskellä syvää epätasapainoa, se ei suhteessa vallitsevaan hintatasoon välttämättä johda suhteessa samansuuruiseen kotimaisten hintojen nousuun – niin kuin devalvaation pitkän ajan neutraalisuutta korostavat, kahden tasapainotilanteen vertailuun perustuvat makroteorian perustulokset väittävät. 1990-luvun vuosikertoja edustaa taas Pentin tähdentämä tarve kiinnittää kokonaistaloudellisissa ennusteissa piste-estimaattien ohella huomiota myös niiden melkoisen suureen luottamusväliin. Ja kyllä moni varmaan muistaa Pentin viime vuosina usein toistaman näkemyksen vapaa-aikaan liittyvän hyödyn merkityksen huomioon ottamisesta.

Pentti Vartian keskustelupuheenvuoroja on aina siivittänyt optimismi. Hän jos kuka on jaksanut nähdä synkätkin tilanteet myönteisessä valossa. Jos kasvu on ollut hidasta, on Pentti jaksanut tähdentää, että kansantalous sentään vielä kasvaa. En suoraan muista häneltä lamavuosien syöksykierteen ajalta sentapaista lausumaa, että laman pohjallakin pystyimme sentään vielä tuottamaan lähes 90 prosenttia lamaa edeltäneen kokonaistuotannon määrästä. Mutta kyllä tuommoinen suhteellisuuden taju olisi hänelle ollut mielestäni tyypillistä.

Ekonomistit ovat niinikään saaneet ammattikuntansa uskollisen puolustajan juuri Pentti Vartiasta. Harvassa taitavat olla ne puheenvuorot, joissa Pentti ei olisi ainakin epäsuorasti tullut korostaneeksi kansantaloudellisen ajattelun käyttökelpoisuutta sekä ekonomistikunnan äänen merkitystä yhteiskunnallisten ilmiöiden erittelyssä. Pentti Vartia, jos kuka, on jaksanut osoittaa kansantaloustieteen perusoppikirjojen käytännön merkitystä.

Eräs hämmentävä piirre Pentissä on se, että samalla kun hän näyttää aina ottavan askeleen vastapuolen suuntaan, hän yhtä säännönmukaisesti pitää lopulta kiinni omasta kannastaan. Joskus hän on tässä suhteessa suoraan sanottuna aika jääräpää. Välistä tuntuu siltä, että Vartian mielipiteet ovat yhtä pysyviä kuin ne perusoppikirjojen keskeiset lauseet, joilla hän niitä perustelee.

Mihinkäs koira karvoistaan pääsisi. Me kaikki kannamme mukanamme omia näkemyksiämme, joita me sovellamme uusiin olosuhteisiin ja joita me ajan myötä tietenkin voimme hioa ja jalostaa. Mutta itse perusnäkemys, visio, pysyy sellaisena kuin millaiseksi se jo nuorena, alle kolmikymppisenä, vakiintui. Tätä korosti aikanaan Joseph Schumpeter monumentaalisessa teoksessaan History of Economic Analysis (1954). Tämän perusnäkemyksen pysyvyys pätee hyvin juuri Pentti Vartiaan. Pohjimmiltaan ekonomisti-Pentti näkee yhteiskunnan – sen rajuista muutoksista huolimatta – perusteiltaan samanlaisena kuin aloittaessaan uransa 1970-luvun alun Katajanokalla.

Sama perusnäkemyksen pysyvyys pätee kieltämättä myös itseeni. Mikään salaisuus ei liioin ole se, että omat yhteiskunnalliset näkemykseni, samoin kuin monet käsitykseni kansantaloustieteestä ja ekonomistin roolistakin, ovat olleet osaksi erilaisia kuin Pentin. Mutta tämä ei ole koskaan estänyt ystävyyttä eikä antoisaa mielipiteiden vaihtoa.

Oman sukupolvemme suomalaisten ekonomistien makrotaloudellisten näkemysten koelaboratorioksi on tullut 1990-luvun lama. Lamatulkintamme paljastaa omien talouspoliittisten näkemystemme ytimen. Siinä ajattelumme makroteoreettiset perusteet joutuvat testiin. Laman aikaisiin ja sitä edeltäneisiin sanomisiimme ja laman tapahtumiin me kaikki joudumme jatkuvasti palaamaan omaa näkemystämme kehittäessämme. Ehkä lama sai ainakin jotkut meistä myös ratkaisevasti muuttamaan omia peruskäsityksiään.

Ekonomistin uransa kypsässä vaiheessa Pentti Vartia valitsi keskeiseksi tutkimuskohteekseen juuri suuren laman. Tämä on luettava hänelle sitä suuremmaksi kunniaksi, kun yhteiset ponnistuksemme laajan, lamaa käsittelevän tutkimushankkeen toteuttamiseksi tuoreeltaan heti laman jälkeen eivät johtaneet rahoittajien penseyden takia muuten kuin osittaisiin tuloksiin. Pentillä riitti intoa jatkaa lamatutkimuksia omana harrastuksenaan. Kun hän sai vielä kumppanikseen aina pystyvän ja aikaansaavan nuoremman kollegan, Jaakko Kianderin, oli tuloksena merkkiteos Suuri lama (1998). Sitä ei kukaan Suomen lamaa käsittelevä tutkija voi sivuuttaa. Vartian ja Kianderin työ on pohjustanut myös pohjoismaista yhteisprojektia, joka aikanaan toivottavasti lisää ymmärrystämme omasta lamastamme.

Vartian ja Kianderin lamakirja on tutkijakuntaa laajemmalle yleisölle tarkoitettu, sujuvasti kirjoitettu yleisesitys lamanaikaisesta ja sitä edeltäneestä taloudellisesta kehityksestä, harjoitetusta talouspolitiikasta sekä tästä käydystä keskustelusta. Kirjan ote on syntetisoiva, eri osapuolten näkökohtia ymmärtävä ja myös yhdistelevä. Sitä ei ole kovin helppo kritisoida.

Eräänlaisen siilipuolustuksen luonnetta kirjassa lisää epäsuorasti se, että sen kirjoittajilla on ollut erilainen käsitys 1990-luvun talouspolitiikan

painotuksista. Jaakko Kiander korostaa enemmän kireän finanssipolitiikan roolia laman syventäjänä, Pentti Vartia taas 1990-luvun alun heikkoa hintakilpailukykyä, vaihtotaseen alijäämää ja liian pitkään jatkunutta markan yliarvostusta. Jaakko siis etsii 1990-luvun laman erityistekijöitä kotimaiselta, Pentti taas avoimelta sektorilta. Tämä "fiskalistien" ja "devalvationistien" vastakkaisuus hallitsee kirjan keskusteluasetelmia. Kirjoittajat suhtautuvat näkemystensä eroihin korrektisti. Mutta tämän molemminpuolisen ymmärtäväisyyden kääntöpuoli on, että näkemysten välinen selvä rajankäynti jää tekemättä.

Laskukauden aikainen kireä finanssipolitiikka ja markan yliarvostus ovat kuuluneet Suomen suhdannekuvaan koko toisen maailmansodan jälkeisen ajan. Devalvationistien ja fiskalistien välinen ristiriita on niinikään ollut jo pitkään nähtävissä – hyvin selvästi ainakin 1970-luvun loppupuoliskolta lähtien. Nämä piirteet eivät tee 1990-luvun lamasta vielä mitään erityistä.

Suuren laman erityispiirre olikin yksityisen sektorin ylivelkaantuminen sekä siihen liittynyt pankkikriisi. Tämä harvinaisen voimakas velkakriisi erottaa suuren laman muista Suomen kansantalouden syvistä laskukausista. Tavanomaiset Suomen laskukausien tunnuspiirteet – valuutan yliarvostus, huono hintakilpailukyky ja laskusuhdannetta kärjistävä, kireä finanssipolitiikka – olivat kyllä kuvassa mukana myös 1990-luvun lamassa. Ne eivät kuitenkaan olleet niin voimakkaita, että ne sinänsä olisivat aiheuttaneet yhtä syvän takaiskun kuin todellisuudessa tapahtui. Epäsuorasti velkakriisin merkittävyydestä kertovat, paitsi samanaikaiset tapahtumat muissa pohjoismaissa, myös myöhemmät pankki- ja valuuttakriisit Kaakkois-Aasiassa ja muualla.

Edellä mainittuja, "fiskalistien" ja "devalvationistien" välisiä käsityseroja itse lamavuosien talouspolitiikasta voi hyvällä syyllä pitää vähempimerkityksisinä kuin niitä tapahtumia, jotka jo aiemmin olivat altistaneet kansantalouden tällaiselle velkakriisille. Kun tukki oli jo päässyt karkaamaan putoukseen, sen ohjastaminen oli mahdotonta. Suomesta oli tullut joksikin aikaa taloudellinen ajopuu.

Suuren laman poikkeuksellisuuden syitä on siis haettava sitä edeltäneestä nousukaudesta. Kriittiset kysymykset kohdistuvat vääjäämättä rahoitusmarkkinoiden vapauttamisen ajoitukseen ja liberalisointitoimien keskinäiseen järjestykseen, niiden vaikutusten huomioon ottamiseen muussa talouspolitiikassa sekä ylipäänsä liberalisoinnille talouspoliittisessa keskustelussa annettuun merkitykseen.

Miten sitten pääoman maahantuonnin paisuminen, luottokannan raju kasvu, asuntojen ja muiden sijoituskohteiden hinnan nousu ja talouden ylikuumeneminen olisi voitu 1980-luvun lopulla estää?

Tarjolla oli kaksi päävaihtoehtoa. Näistä ensimmäisessä rahapolitiikan teho olisi pyritty säilyttämään uusissakin olosuhteissa. Kuten Vartia ja Kianderkin totetavat, rahoitusmarkkinoiden säännöstelyn ja pääomanliikkeiden vapauttamisen viime vaihetta olisi voitu 1980-luvun loppupuoliskolla lykätä, kun nousukauden voima ja rahoitusmarkkinoiden ylikuumeneminen oli käynyt ilmeiseksi. Säännöstelytoimia olisi voitu ehkä tilapäisesti kiristääkin.

Vielä tehokkaampi keino rahapolitiikan tehon palauttamiseksi olisi ehkä ollut markan päästäminen kellumaan viimeisten merkittävien liberalisointitoimien yhteydessä 1980-luvun puolivälissä jo ennen nousukauden voimistumista. Onhan jo oppikirjojen perusasetelmia, että rahapoliittisen autonomian säilyttäminen kansainvälisten pääomanliikkeiden ollessa vapaita edellyttää valuutan päästämistä kellumaan.

Toinen periaatteellinen päävaihtoehto, johon meillä käytännössä lähinnä päädyttiin, oli alun perin omaksutun rahoitusmarkkinoiden asteittaisen vapauttamisen jatkaminen sekä sen seurauksena rahapoliittisesta autonomiasta luopuminen ja sitoutuminen kiinteään valuuttakurssiin. Paitsi finanssipolitiikan erittäin voimakasta kiristämistä vuosikymmenen lopun nousukauden aikana, olisi tämä edellyttänyt kansantalouden vakauttamiseksi vuorostaan sitä, että jollakin keinoin, esimerkiksi verotuksellisin muutoksin ja rahoitusmarkkinoiden valvonnan tehostamisen avulla, yleisö ja pankit olisi saatu tiedostamaan uuden rahoitusympäristön riskit ja varautumaan niihin paremmin.

Miten sitten suomalaiset ekonomistit suhtautuivat 1980-luvulla näihin vaihtoehtoihin? Miten he pyrkivät vakuuttamaan poliittiset päätöksentekijät esittämiensä toimenpiteiden välttämättömyydestä?

Suomalaisessa ekonomistikunnassa ei tuolloin ollut edustettuna selvästi jäsentynyttä mielipidettä, joka olisi johdonmukaisesti ja konkreettisesti esittänyt toimia rahoitusmarkkinoiden säännöstelyn aikataulun ja marssijärjestyksen muuttamiseksi tai pääomanliikkeiden vapauttamistoimien lykkäämiseksi sopivampaan ajankohtaan. Yksittäisiä puheenvuoroja oli – itsekin muistan niitä aikoinaan esittäneeni – mutta nämä jäivät painoarvoltaan heikoiksi ja epäkonkreettisiksi. Näitä puheenvuoroja esittäneet epäilevät tuomaat taisivat itsekin ajatella, että karavaani joka tapauksessa kulkee heidän epäilyistään huolimatta. He tunsivat itsensä vanhanaikaisiksi peräpeiliin tuijottajiksi, joiden puheet oli jo ennakolta leimattu turhiksi.

Sama tuonpuoleisen maailman sävy liittyi 1980-luvun lopulla myös kelluvan valuuttakurssin vaihtoehtoon. Oppikirjatapauksena se kyllä tunnettiin. Mutta ekonomistiemme keskuudessa ajateltiin yleisesti, että tämä vaihtoehto oli epäkelpo pieneen, tuotantorakenteeltaan yksipuoliseen ja inflaatioherkkään kansantalouteemme. Australian ja Uuden Seelannin huonot kokemukset kellutuksesta tarjosivat tässä suhteessa varoittavan esimerkin. Ne ekonomistimme, jotka tuolloin puhuivat kellutuksesta myönteiseen sävyyn, kuten esimerkiksi Björn Wahlroos, tarjosivat tätä vaihtoehtoa lähinnä jonkinlaisena hayekilaisena vapaan rahan valtakuntana, josta julkisen vallan sääntely olisi ylipäänsä kaikonnut. Myöskään Vartia ja Kiander eivät kuitenkaan kirjassaan paljoakaan käsittele tätä vaihtoehtoa. He keskittyvät keskustelemaan lähinnä kertaluonteisesta devalvaatiosta kiinteän valuuttakurssijärjestelmän puitteissa.

Suomalainen ekonomistikunta oli yleensäkin ankkuroitunut ajattelussaan kiinteään valuuttakurssiin 1990-luvun alkuun saakka. Talouspolitiikkaa valmistelevien ekonomistien piirissä erityisesti Suomen Pankissa peruuttamattomasti kiinteätä valuuttakurssia pidettiin sinä ankkurina, johon sitoutumalla devalvaatiosyklistä voitaisiin irrottautua ja inflaatio saataisiin hillintään. Kilpailijamaita nopeamman inflaation aiheuttama työttömyys nähtiin kyllä tämän vaihtoehdon riskinä. Mutta ajateltiin, että mitä paremmin työmarkkinaosapuolet tiedostavat korkeiden palkankorotusten, inflaation ja työttömyyden välisen yhteyden, sitä paremmin tämä riski voidaan välttää. Ei suljettu pois sitäkään, että jonkinlainen työttömyyden kasvun muodossa tuleva läksy olisi tarpeen tämän kiinteiden valuuttakurssien tapauksessa vääjäämättömän yhteyden sisäistämiseksi.

Nämä kiinteän valuuttakurssin tuomaa kurinalaisuutta korostavat ekonomistimme pitivät välttämättömänä myös ulkomaisten pääomanliikkeiden ja kotimaisten rahoitusmarkkinoiden säännöstelyn purkamista. Tämä oli tietysti tarpeellista jo resurssien tehokkaan allokaation kannalta. Rahoitusmarkkinoiden säännöstelyn purku myös selvensi finanssi- ja rahapolitiikan välisen työnjaon. Tämähän kiteytyi kohta kuuluisaksi työnjako-opiksi, joka sysäsi vastuun työllisyydestä finanssipolitiikan ja työmarkkinajärjestöjen vastuulle. Ulkomaiset pääomanliikkeet olivat lisäksi omiaan tehostamaan markkinakuria. Pääoman vienti saattoi joissakin olosuhteissa voimistaa esimerkiksi inflaation kiihtymisen vaikutusta työttömyyden kasvuun ajamalla kotimaisia korkoja ylöspäin.

Miksi tämä rahataloudellisen vakauden eksperimentti sortui pankkikriisiin ja lamaan? Rahoitusmarkkinoitamme läheltä seuranneet ja niiden säätelyä koskevia päätöksiä valmistelleet ekonomistimme ovat olleet lähinnä

sitä mieltä, että pääsyyt löytyvät itse säännöstelyn kaudelta (tuoreimpana esimerkkinä vrt. esim. Suvanto, 2002). Säännöstelykauden aikana sekä pankkien, yleisön että poliittisten päätöksentekijöiden ajatustottumuksiin oli pesiytynyt vääristymiä, jotka säännöstelyn purkauduttua vääjäämättä johtivat katastrofiin. Säännöstely purettiin Suomesta kerta kaikkiaan liian myöhään. Sisäisen järjestelmämme ja käyttäytymistottumustemme ristiriita kansainvälisen toimintaympäristön kanssa oli jo kärjistynyt pitelemättömäksi.

Nämä asioita läheltä seuranneet ekonomistit olivat tietoisia siitä, että rahoitusmarkkinoiden vapauttaminen olisi edellyttänyt säännöstelyn kaudelta periytyneiden verotuksellisten vääristymien korjaamista. Nämä vääristymät kannustivat ylenmääräiseen velkaantumiseen, vähensivät pankkien välistä kilpailua ja aiheuttivat tehottomuutta pääomien käytössä. Pankkitalletusten verovapaus olisi täytynyt purkaa ja supistaa korkojen vähennysoikeuksia verotuksessa. Rahoituslaitosten riskien valvonnan tarve tiedostettiin myös ainakin periaatteellisella tasolla, samoin vakavaraisuussäädösten jälkeenjääneisyys ja tehottomuus.

Mutta ongelma oli, että säännöstelyn rippeisiin liittyi etunäkökohtia eivätkä poliittiset päätöksentekijät olleet valmiita ryhtymään vääristymiä korjaaviin toimiin. Eduskunnassa oli vahva "pankkipuolue", joka torjui esimerkiksi talletusten verovapauden ja siihen liittyvän talletuskorkokartellin purkamisen. Lainakorkojen verovähennysoikeuksien supistamista vastustettiin niinikään. Pankkitarkastus oli vallitsevin säädöksin voimaton, ja pankit vehkeilivät vakavaraisuussäädösten sekä rahoitustarkastuksen kiristämistä vastaan. Säännöstelyjärjestelmästä periytyneet intressit ja ajatustottumukset olivat kääriytyneet yhdeksi, aukeamattomaksi keräksi.

Tämän virkamiestulkinnan mukaan siis vanha, mätä järjestelmä eteni vääjäämättä kohti tuhoaan. Ulkopuoliset näkivät katastrofin merkit ja varoittivatkin niistä. Mutta tämä oli turhaa. Vika vain ei ollut virkamiesten vaan päätöksentekijöiden.

Tällainen jälkikäteisarvio paljolti sivuuttaa kysymyksen siitä, millainen "policy design" olisi ollut tarpeen liberalisointiin liittyvän siirtymävaiheen hoitamiseksi mahdollisimman vähäisin kolhuin. Kysymys asetetaan koskemaan lähinnä vain sitä, kumpi järjestelmä, vanha säännöstely vai liberalisoitu, on parempi, tehokkaampi ja lopulta elinkykyisempi. Mutta todellisuudessa kysymys ei 1980-luvun loppuvuosina ollut vain tästä. Kun lähtökohtana kerran oli säännöstellyt markkinat, edellytti liberalisointi välttämättä siirtymävaihetta, johon liittyivät omat, vaikeat ongelmansa. Säännöstelyn pahasta perimästä ei päässyt eroon vain toteamalla, että se oli poliitikkojen ja pankki-

puolueen vika. Se varjosti liberalisointivaihetta, ja sen vaikutukset olisi täytynyt ottaa tätä suunniteltaessa huomioon.

Millaisen loppulausunnon Vartia ja Kiander sitten antavat oman ammattikuntansa toiminnasta noina kriittisinä vuosina? Luettuani heidän kirjansa uudelleen minulle jäi käsitys, että myös he päätyvät antamaan ekonomisteille pitkälti samankaltaisen synninpäästön kuin edellä esitetty virkamiestulkintakin. Ongelma oli poliitikkojen ja muiden päätöksentekijöiden kuurous, ei asiantuntijoiden puutteellinen käsitys siitä, mitä oli tapahtumassa.

Vartian ja Kianderin arvio on tosin astetta nyansoidumpi. Tekijät toteavat, että liberalisointia ei ekonomistien toimesta etukäteen tutkittu riittävästi, ei Suomen Pankin toimesta eikä muutenkaan. Siitä ei liioin käyty keskustelua. Puheenvuorot jäivät yksittäisiksi. Tämä on totta.

Vartia ja Kiander antavat toisaalta ymmärtää, että asiantuntijoilla oli yhtä kaikki haju siitä, mitä oli tulossa. Esimerkiksi Pertti Kukkosen vuonna 1984 käyttämä puheenvuoro säännöstelyn purkamisen riskeistä noteerataan, samoin Matti Rankin eduskunnassa syksyllä 1986 esittämä, suorastaan profetiaksi luonnehdittu pohdiskelu luottobuumia seuraavasta haaksirikosta. Pannaan myös merkille, että Suomen Pankissa tehtiin useita liberalisointia koskevia selvityksiä, julkaisemattomia muistioita ilmeisesti vielä enemmän.

Vartia ja Kiander korostavat, että poliittiset päätöksentekijät olivat kuuroja näille ekonomistien kaukonäköisille varoituksille. He panevat aivan oikein merkille, että ekonomistit vaativat rahoitusmarkkinoiden vapauttamisen yhteydessä esimerkiksi talletusten verovapaudesta luopumista ja talletuskorkokartellin kumoamista sekä lainakorkojen verovähennysoikeuden rajoittamista. Nämä vaatimukset kilpistyivät kuitenkin poliittisen tahon vastustukseen. "Poliitikkojen oli vaikea viedä läpi ... verojen kiristämistä tai vähennysoikeuksien kaventamista vastoin äänestäjien tahtoa." (Kiander – Vartia 1998, 287)

Vielä rahoitusmarkkinoiden vapauttamista enemmän ekonomistit olivat huolissaan suhdannepolitiikan kokonaisvirityksen keveydestä. Vartia ja Kiander niinikään korostavat ekonomistiemme ymmärtäneen 1980-luvun lopulla hyvin, että kiinteä valuuttakurssi merkitsee rahapolitiikan tehon ehtymistä ulkomaisten pääomanliikkeiden vapauduttua. Vuosikymmenen lopun yllättävän nousukauden aikana finanssipolitiikkaa olisi siten täytynyt kiristää voimakkaasti. Tätä ekonomistit ja Suomen Pankki esittivätkin ja tähän myös hallitus pyrki, mutta kireän linjan läpivieminen osoittautui poliittisesti mahdottomaksi. Vika oli siis tässäkin poliitikkojen.

Tämä kaikki on totta. Me 1980-luvun lopun talouspoliittisessa keskustelussa mukana olleet muistamme kaikki hyvin pankkipuolueen mahdin talletuskorkokartellin ylläpitämisessä samoin kuin poliittisten päätöksentekijöiden haluttomuuden finanssipolitiikan riittävään kiristämiseen. Poliitikot olivat enemmän huolissaan veronalennusten riittämättömyydestä kuin finanssipolitiikan löysyydestä.

Mutta voimmeko siltikään puhtaalla omallatunnolla sysätä syytä yksinomaan lyhytnäköisten ja taktisesti suhtautuvien poliittisten päätöksentekijöiden niskaan? Eikö vastuuntuntoisten ekonomistien rooliin olisi sitä suuremmalla syyllä kuulunut toimia päätöksentekijöiden herättäjänä? Mitä suuremmat intressit olivat kietoutuneet yhteen säännöstelyn rippeiden ja löysän finanssipolitiikan kanssa, sitä äänekkäämmin olisi hälytyskelloja kuulunut soittaa.

Eikö rehellisyyden nimissä olisi myönnettävä, että meillä oli itsellämme etenkin rahoitusmarkkinoiden vapauttamisen merkityksestä ja riskeistä niin hatara kuva, että meillä ei ollut pontta eikä rohkeutta päätöksentekijöiden ravisteluun? Mekään emme tajunneet, millainen aikapommi kansantalouteen tuli 1980-luvun lopun luottobuumin aikana viritetyksi? Tämä ymmärryksen puute vei ääneltämme voiman.

Tämän seikan myöntämistä helpottaa se, että velkakriisit olivat 1980-luvun lopun oloissa tosiaan lähinnä utuista historiaa. Olimme voineet tutustua niihin lähinnä vain taloushistoriallisten kuvausten perustella. Noissa oloissa erehtyminen oli perin inhimillistä.

Vartia ja Kiander osoittavat mm. Kukkosen ja Rankin esimerkeillä, että esimerkiksi liberalisoinnin epäilijöitä löytyi. Mutta tämä ei ole pääasia. Ainahan kirjoitellaan kaikenlaista. Pääasia on, oliko epäilijöiden ääni riittävän vakuuttava ja oliko heillä tarjota sellaista konkreettista vaihtoehtoa, että 1980-lopun talouspoliittinen kokonaistilanne olisi joutunut perusteelliseen läpivalaisuun. Tässäkään asiassa eivät yksittäiset pääskyt riittäneet tekemään kesää. Erillisistä puheenvuoroista ei muodostunut sellaista jäsenneltyä kokonaisuutta, joka olisi pakottanut muut ottamaan kantaa. Poikkeavat puheenvuorot on kaivettu esiin vasta jälkikäteen, kun kriisin taustoja on ruvettu penkomaan. Omana aikanaan ne eivät kantaneet.

On totta, että esimerkiksi säännöstelyjärjestelmään liittyneillä verotuksellisilla vääristymillä oli vahva poliittinen kannatus. Näinhän laajoja väestöpiirejä suosivilla veroeduilla aina on. Mutta juuri nämä vanhaan järjestelmään sisäänrakennetut edut olisivat tehneet tarpeelliseksi sen, että rahoitusmarkkinoiden vapauttamisesta olisi käyty perinpohjainen yhteiskunnallinen keskustelu. Tämän keskustelun ja sitä seuraavan päätöksenteon pohjaksi asi-

antuntijoiden olisi täytynyt valmistella esimerkiksi komiteatyönä perusteellinen aineisto. Tässä olisi voitu tehdä arvio liberalisoinnin eduista ja haitoista, liberalisoinnin yhteydessä noudatettavasta strategiasta sekä liberalisoinnin edellyttämistä verolainsäädännön muutoksista sekä sen jälkeen noudatettavasta rahapolitiikan strategiasta samoin kuin tämän vaikutuksista muuhun talouspolitiikkaan.

Kun rahoitusmarkkinoiden liberalisointi olisi nostettu tällä tavoin – tosiasiallisen merkityksensä mukaisesti – keskeiseksi yhteiskunnalliseksi kysymykseksi, yleisö ja poliittiset päätöksentekijät olisivat voineet tiedostaa paremmin meneillään olevan muutoksen tärkeyden. Asiaan olisi paneuduttu, vastuuntunto olisi lisääntynyt ja uuden ympäristön riskeihin olisi varauduttu paremmin sekä pankkien, yleisön että talouspoliittisten päätöksentekijöiden keskuudessa.

Rahoitusmarkkinoiden muutos oli yhteiskunnalliselta merkitykseltään niin suuri asia, että se olisi toisin sanoen pitänyt aidosti politisoida. Myös finanssipolitiikan uusi, merkittävämpi rooli olisi voitu nostaa esiin samassa yhteydessä. Finanssipolitiikan suhdanneluonteinen kiristämisen tarve olisi tullut paremmin perustelluksi.

Mutta talouspoliittinen valmistelukoneisto tuskin lopultakaan halusi tällaista rahoitusmarkkinoiden järjestelmämuutoksen politisointia. Asteittainen liberalisointi oli jo aiemmin totuttu viemään läpiin epäpoliittisina, hallinnollisina rutiinipäätöksinä. Poliittiset päätöksentekijät oli pidetty siitä sivussa. Ajoittaisilla väliintuloillaan nämä olivatkin aiheuttaneet lähinnä vain harmia. Mutta sivuun jääminen vähensi toisaalta poliittisen tahon vastuuntuntoa ja lisäsi houkutusta poliittisten irtopisteiden kalasteluun.

Tämän perinteisen asiantuntijavetoisuuden ohella tarvetta järjestelmämuutoksen nostamiseksi poliittisen päätöksenteon agendalle oli omiaan vähentämään se, että säännöstelyn purkamista pidettiin hyvänä asiana sinänsä. Sen ei haluttu sortuvan toisarvoisiin poliittisiin kiistoihin. Siirtymävaiheen ongelmia ei myöskään pidetty vakavina. Asiantuntijat eivät nähneet siihen liittyviä riskejä.

Talouspolitiikan valmistelijoiden päämielenkiinto kohdistui 1980-luvun lopulla sitä paitsi inflaatiokurin ankkuroimiseen peruuttamattomasti kiinteän valuuttakurssin avulla. Kun finanssipolitiikan kiristämisessä tässä ei onnistuttu kovin hyvin, pidettiin pääomanliikkeiden vapauttamisen kautta tapahtuvan markkinakurin tehostumista hyvänä asiana.

Suomalaiselle talouspoliittiselle perinteelle on virkamiesvetoisuuteen liittyen yleensäkin ollut ominaista eräänlainen olosuhteiden painon hyväksi käyttäminen etukäteisen valmistautumisen, keskustelun ja suunnittelun sijas-

ta. Näin meillä 1950-luvun lopulta lähtien paljolti vastaanotettiin vapaakauppa, rakennemuutos ja maaltamuutto. Rahoitusmarkkinoiden vapauttaminen edusti 1980-luvulla samaa perinnettä. Usein tällainen hyppy tuntemattomaan onnistui hyvin. Mutta liberalisoinnin loppuvaiheissa se johti katastrofiin.

Jonkinlainen menettelytapojen muutos näyttää olevan meneillään. Esimerkiksi 1990-luvulla toteutettujen EU- ja EMU-ratkaisujen yhteydessä käytiin perusteellisempi ja laaja-alaisempi keskustelu- ja valmistautumiskierros. Tämän myötä tietämys uusien olosuhteiden asettamista haasteista lisääntyi. Lisäksi toteutettiin uudistuksia, joilla pyrittiin helpottamaan uuteen ympäristöön sopeutumista.

Emme voi tietää, olisiko edes huolellisemmin suunniteltu, poliittisesti vahvistettu liberalisointistrategia estänyt pankkikriisin. Se, että vastaavankaltaisia kriisejä puhkesi muutama vuosi myöhemmin eri puolilla maailmaa ja ne yllättivät myös vahvoin asiantuntijavoimin varustetun IMF:n, puhuu sen puolesta, että kriisiä tuskin olisi voitu meilläkään kokonaan välttää. Suomen kriisi sitä paitsi puhkesi suhteellisen varhain. Tuoreita kokemuksia vastaavista kriiseistä oli vähän.

Tuskin olemme päässeet rahoituskriiseistä vieläkään eroon. Omien 1990-luvun alun kokemustemme näkökulmasta on pelottavaa ajatella, miten pitkälle vietyä kotitalouksien ja yritysten velkaantuminen sekä asuntomarkkinoiden ylikuumeneminen on tällä hetkellä Yhdysvaltojen ja Englannin kaltaisissa maissa. Velkadeflaation puhkeaminen Yhdysvalloissa voisi laukaista maailmanlaajuisen laman. Suomen kaltaisessa pienessä maassa tapahtuvalla valuutan arvon heikkenemisellä ja vaihtotaseen vahvistumisella on vähäinen vaikutus muuhun maailmaan. Yhdysvaltojen tapauksessa kontraktiiviset vaikutukset muihin maihin olisivat kokonaan toista mittaluokkaa.

Suomi käy hyvin esimerkiksi siitä, että läntisessä teollisuusmaassa laman varjo voi olla, kuten meillä on totuttu sanomaan, pitkä. Se on jopa pidempi kuin Vartia ja Kiander antavat ymmärtää. Heijastaneekohan Pentti Vartian valoisaa elämänasennetta se, että tekijät kirjansa lopuksi tulevat varsin pitkälle suhteellistaneeksi koko laman: "Lama tulee ennen pitkää näyttäytymään vain väliaikaisena pudotuksena kansantalouden pitkällä kasvuuralla"(Kiander – Vartia 1998, 324), ja siitä on ollut hyötyäkin, koska sen "...aiheuttama 'luova tuho' on nostanut esiin uusia kasvavia yrityksiä ja toimialoja" (Kiander – Vartia 1998, 325). Totta kyllä, bruttokansantuote saavutti jo 1990-luvun loppuun mennessä lähes lamaa edeltäneeltä kaudelta vedetyn trendin mukaisen tasonsa. Kansantalouden rakenteellinen uusiutuminen on ollut niinikään voimakasta. Viennin osuus bruttokansantuotteesta on 1980-

luvun lukemista liki kaksinkertaistunut, ja sähköteknisestä teollisuudesta on tullut paperin rinnalle tärkein vientiala.

Mutta myös laman kielteiset seuraukset näkyvät yhä. Kymmenillä tuhansilla laman aikana työpaikkansa menettäneillä tai työmarkkinoille tuolloin tulleilla on yhä vaikeuksia löytää paikkansa työelämässä. Rakenteellinen työttömyys kohosi laman jälkeen kokonaan uudelle tasolle. Konkurssien myötä poistui tuotantoa, joka olisi ollut pidemmän päälle kilpailukykyistä ja kannattavaa. Tällaiset laman pysyväisluonteiset vaikutukset tekevät entistä tärkeämmäksi vastaisen varalle oppia hakevan itsetutkiskelun ekonomistien roolista laman merkkien tulkitsijana.

Vartia ja Kiander toteavat kirjansa lopuksi aivan oikein, että yllättävien kriisien mahdollisuus on kansantaloudessa aina olemassa. Niihin tulisi varautua esimerkiksi vaihtoehtoisten tulevaisuudenkuvien ennakkoluulottomalla pohdiskelulla. Keskustelun kirjoon tulisi aina sisältyä myös meneillään olevien rakenteellisten muutosten outojen ja epätodennäköisten seurausten arviointi. Talouden muutos voi olla niin kaleidoskooppimaista, murrosmaisesti uudelleen rakentuvaa, että vakiintuneeseen kehitykseen perustuvat tulevaisuudenskenaariot pitäisi aina uskaltaa alistaa kriittiseen syyniin.

Tähän arvaamattomuuteen varautuminen edellyttää sitä, että myös kylähulluina pidettyjen ääntä tulisi kuulla. Erityisesti Suomessa, jossa on totuttu korostamaan yksimielisyyden ja samassa veneessä istumisen tarvetta, meidän tulisi muistaa tämä henkisen avarakatseisuuden tarve. Se on eräs tapa varautua epävarmaan tulevaisuuteen. Vartian ja Kianderin teos antaa hyvän oppitunnin siitä, millaisiksi voivat muodostua seuraukset, jos yhteiskunta päästää henkiset tuntosarvensa ruostumaan. Jo yksin tästä syystä kannattaa ystävieni merkkiteosta lukea ja kommentoida vielä vuosienkin päästä.

KIRJALLISUUTTA

Kiander, Jaakko ja Vartia, Pentti (1998): Suuri lama. Suomen 1990-luvun kriisi ja talouspoliittinen keskustelu. Taloustieto Oy, Helsinki.

Schumpeter, Joseph A. (1954): History of Economic Analysis, Oxford University Press, Lontoo.

Suvanto, Antti (2002): Miten Suomi muuttui?, teoksessa Kajanoja, Jouko, Kiander, Jaakko, Voipio, Iikko B. (toim.): Tutkimus ja talouspolitiikka. Kirjoituksia taloudellisesta kasvusta.Valtion taloudellinen tutkimuskeskus, Helsinki.