

UNITED
BY OUR
DIFFERENCE



PROMEMORIA

Validering av Mocca

2013-12-16

Analys & Strategi

Konsulter inom samhällsutveckling

WSP Analys & Strategi är en konsultverksamhet inom samhällsutveckling. Vi arbetar på uppdrag av myndigheter, företag och organisationer för att bidra till ett samhälle anpassat för samtiden såväl som framtiden. Vi förstår de utmaningar som våra uppdragsgivare ställs inför, och bistår med kunskap som hjälper dem hantera det komplexa förhållandet mellan människor, natur och byggd miljö.

Innehåll

INLEDNING	5
Syfte	5
Mocca	5
VALIDERINGSMETODIK	7
DATA.....	8
RESULTAT.....	9
Faktisk utvecklingen av löner, bostadspriser och pendling	9
Validering.....	10
Omestimering av Mocca	15
Övrigt arbete.....	16
Orsaker till den höga känsligheten för löneskillnader	19

Inledning

Syfte

Mocca är en dynamisk modell för arbetspendlingen över Öresund med tillgänglighet, bostadspris- och löneskillnader som prognosvariabler. Att modellen är dynamisk innebär att den ger successiva prognoser år för år från ett utgångsläge till ett givet prognosår. Mocca har använts som en delmodell för att prognostisera effekter av infrastrukturåtgärder i Öresundsregionen.

Mocca utvecklades på data för åren 2001-2006. Nu finns data av samma slag tillgängliga för perioden 2007-2011. Detta ger möjlighet att validera hur Mocca presterar som prognosmodell i en så kallad ”out-of-sample” validering. Syftet med att innevarande forskningsprojektet är att utföra en sådan validering av Mocca. Ytterligare ett syfte är att studera möjligheterna att flytta Mocca för prognoser i andra gränsregioner än Öresund.

Den ursprungliga estimeringen av Mocca utfördes på enbart en delmängd av det då tillgängliga datamaterialet. Mocca har under projektet estimerats om på det fullständiga datamaterialet. Utöver detta så har lämpliga sätt att kalibrera Mocca studerats.

Projektet har genomförts av Christer Persson och Karin Brundell-Freij, WSP, på uppdrag av Trafikverket.

Mocca

Mocca är som sagts ovan en dynamisk modell för arbetspendlingen över Öresund med tillgänglighet, bostadspris- och löneskillnader som prognosvariabler. Regionen som igår i Mocca är den så kallade Öresundsregionen bestående av Skåne och Själland. Mocca modellerar individens dynamiska val av bostads- och arbetsplatsområden, vilket ger en prognos för arbetspendlingen över Öresund. Denna del av modellen är uppbyggd som en så kallad Markovkedja, vilken bestämmer dynamiken i Mocca. Den viktigaste komponenten i en Markovkedja är övergångsmatrisen som anger individens övergångssannolikheter mellan par av bostads- och arbetsplatsområden från ett år till nästföljande år. Övergångssannolikheterna modelleras som en logitmodell där tre prognosvariabler ingår:

- Genomsnittlig löneskillnad mellan danska och svenska sidan i regionen
- Genomsnittliga bostadsprisskillnader för enfamiljshus mellan den danska och svenska sidan i regionen
- Tillgänglighet i form av generaliserad transportkostnad mellan par av bostads- och arbetsplatsområden

Givet en ursprunglig fördelning av individerna över bostads- och arbetsplatsområden för ett utgångsår så beräknar Mocca successivt med hjälp av övergångssannolikheterna individernas fördelning över bostads- och arbetsplatsområden år för år.

Tillgängligheten beräknas med hjälp av en underliggande transportmodell för arbetsresandet över Öresund. Transportmodellen är en traditionell logitmodell för val av korsningspunkt och transportmedel över Öresund. Utöver att beräkna generaliserad transportkostnad som går in till modellen för övergångssannolikheterna, så fördelar transportmodellen Öresundspendlarnas arbetsresor på korsningspunkt och transportmedel.

De två delmodellerna, övergångssannolikheterna och transportmodellen, utgör tillsammans en fyrstegs transportmodell för arbetsresor. Denna modell har en utmärkande egenskap som särskiljer den från en traditionell fyrstegsmodell, nämligen:

- Givet en ursprunglig fördelning av individerna över bostads- och arbetsplatsområden så är valet av bostadsområde endogeniserat i modellen för efterföljande år

Att valet av bostadsområde är endogeniserat innebär att det bestäms av modellen och kan inte användas som prognosvariabel. Det fjärde steget, val av rutt i transportnätet, är rudimentärt och består enbart av val av korsningspunkt över Öresund, och är som nämnts ovan kombinerat med val av transportmedel.

En Markovkedja av det slag som ingår i Mocca har den egenskapen att givet tillräckligt lång tid når den godtyckligt nära ett så kallat stationärt tillstånd där fördelningen över tillstånden (bostads- och arbetsplatsområden) inte ändrar sig från år till år. I den meningen ger Mocca samtidigt den dynamiska förändringen av arbetspendlingen från år till år, samtidigt som modellen ger ett långsiktigt jämviktssläge för arbetspendlingen.

För en mer detaljerad beskrivning av Mocca hänvisas till rapporten *MOCCA - modell för pendlingen över Öresund. Modelluppbyggnad och manual* (WSP 2009-11-30).

Valideringsmetodik

Utgångspunkten för projektet är att utvärdera Moccas prediktionsförmåga för data som modellen inte är estimerad på, alltså data efter 2006 om arbetspendling över Öresund. Data av samma slag som Mocca estimerades på finns tillgängligt fram till och med år 2010/2011. Moccas prognoser över arbetsresandet bygger på en trafikmodell som estimerades på data från en resvaneundersökning om resandet över Öresund som genomfördes 2009. Ingen ny sådan resvaneundersökning har genomförts sedan dess, därför har valideringen fokuserat på Moccas prognoser av antalet årliga arbetspendlare över Öresund. Antalet pendlare sätter dock nivån för arbetsresorna och är därmed i hög grad styrande för Mocca prognosförmåga för arbetsresorna.

Det finns ett flertal formella testmetoder för att avgöra en modells prediktiva förmåga, dessa ger p-värden eller signifikansnivåerna för hur väl modellens prognoser överensstämmer med verkliga data. De avvikelser som finns mellan Mocca och verkliga data är dock av den storleksordningen att formella statistiska test inte adderar någon verklig information till resultaten. Därför presenteras valideringsresultaten, i denna promemoria, enbart i diagramform.

Utgångspunkten är att använda de verkliga värdena på prognosvariablerna för valideringsperioden i prognoskörningar med Mocca, och jämföra prognosresultaten för arbetspendlingen över Öresund med verkliga data över pendlingen.

Den fullständiga uppsättningen av valideringsdata för löner, bostadspriser och pendlare finns fram till och med år 2010. Under perioden 2006-2010 skedde enbart marginella förändringar i transportsystemet avseende resor över Öresund. Den största förändringen var bygget av citytunneln i Malmö. Den öppnades dock för trafik 12 december 2010 och har därmed i praktiken inte haft någon påverkan på resultaten. I prognoskörningarna har tillgängligheten hanterats som oförändrad under hela valideringsperioden 2006-2010. Därmed är det också givet att valideringen gäller hur väl Mocca kan prognostisera förändringar i pendlingen som beror på förändrade löner och bostadspriser i regionen.

Data

De data som har använts för validering kommer från samma källa som användes vid den ursprungliga estimeringen av Mocca, det dansk-svenska statistiksamarbetet Örestat¹. Uppgifterna har hämtats ur följande tabeller:

- Löner, nettolön efter skatt: Tabell *Inc 2b, Kolumn K – Personal income after tax*
- Bostadspriser, genomsnittliga försäljningspris för enfamiljshus: Tabell *Bo13, Average price Single-family houses*
- Arbetspendling: från rapporten *Tendens Öresund 2012* (del av dokumentation från Örestat), tabell på sidan 65.

För löner och bostadspriser används den absoluta skillnaden mellan respektive uppgift för

- region Öresund DK
- region Öresund SE

mätt i 1000-tals svenska kronor. Värdena är typiskt >0, vilket betyder att både löner och bostadspriser är högre i Danmark än i Sverige.

För arbetspendlingen används antalet personer per år som arbetspendlar.

Tabellen nedan visar utvecklingen för dessa tre dataserier mellan åren 2001 och 2010.

Tabell 1: Arbetspendling, bostadsprisskillnader och löneskillnader faktisk utveckling 2001-2010 (källa: Örestat).

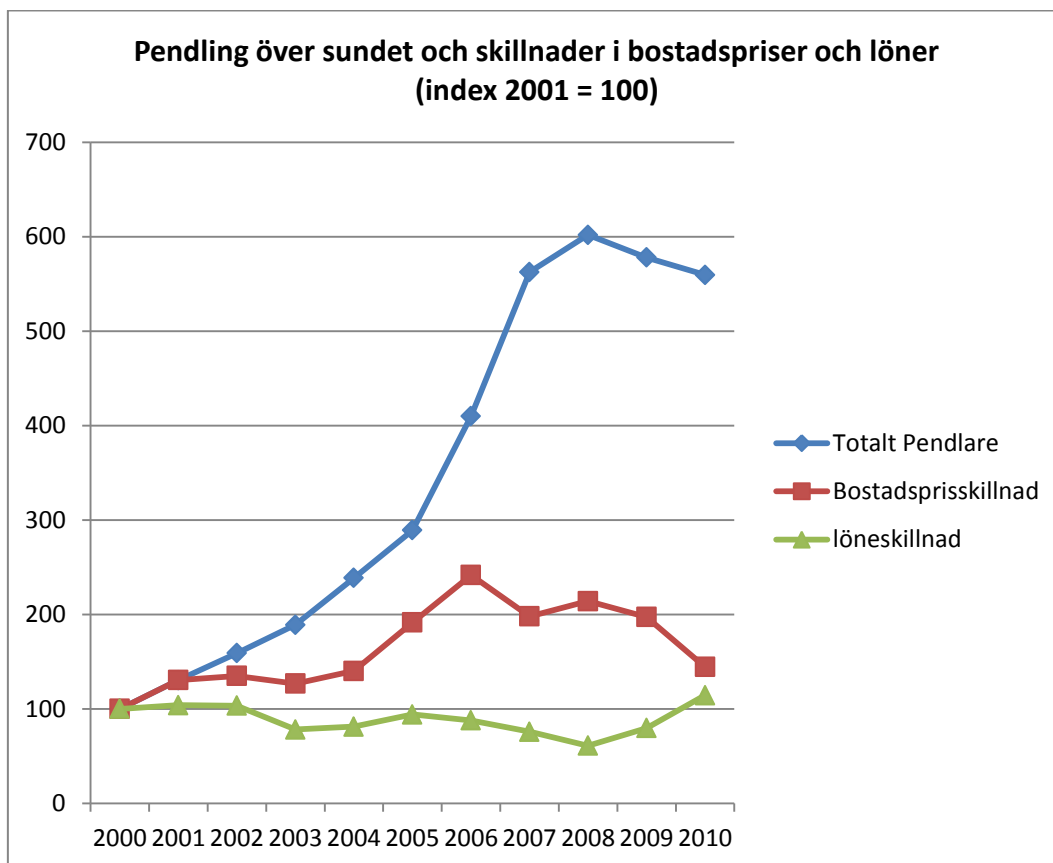
År	Pendlare	Bostadsprisskillnad (1000 kr)	Löneskillnad (1000 kr)
2000	3 291	565	32,2
2001	4 290	737	33,5
2002	5 230	762	33,3
2003	6 219	716	25,1
2004	7 852	792	26,1
2005	9 517	1 081	30,3
2006	13 494	1 365	30,2
2007	18 513	1 118	24,4
2008	19 805	1 208	19,7
2009	19 020	1 114	25,7
2010	18 414	815	36,9

¹ www.orestat.scb.se

Resultat

Faktisk utvecklingen av löner, bostadspriser och pendling

Figur 1 nedan visar samma data som tabell 1 ovan, men för att åskådligt få in alla tre dataserierna i samma diagram så har de räknats om till indexserier med år 2001 som basår.



Figur 1: Arbetspendling, bostadsprisskillnader och löneskillnader faktisk utveckling (index) 2001-2010 (källa: Örestat).

Arbetspendlingen har alltså ökat sexfaldigt under perioden. Bostadsprisskillnaderna mellan danska och svenska sidan mer än fördubblades fram till 2006, för att sedan falla tillbaka. För löneskillnaderna finns ingen klar tendens, men det går att se dem som ungefär densamma över hela perioden

Om Mocca är en korrekt modell för att beskriva arbetspendlingen över sundet så beror pendlingsökningen till allra största delen av tillgänglighetsprånget som skedde i och med öppnandet av Öresundsbron 1 juli 2000, men att ökning har förstärkts av de ökade skillnaderna i bostadspris. Att tillgänglighetsprånget inte direkt leder till ökad pendling utan att det är en utveckling under en följd av år beror (enligt Mocca) på att det tar tid för in-

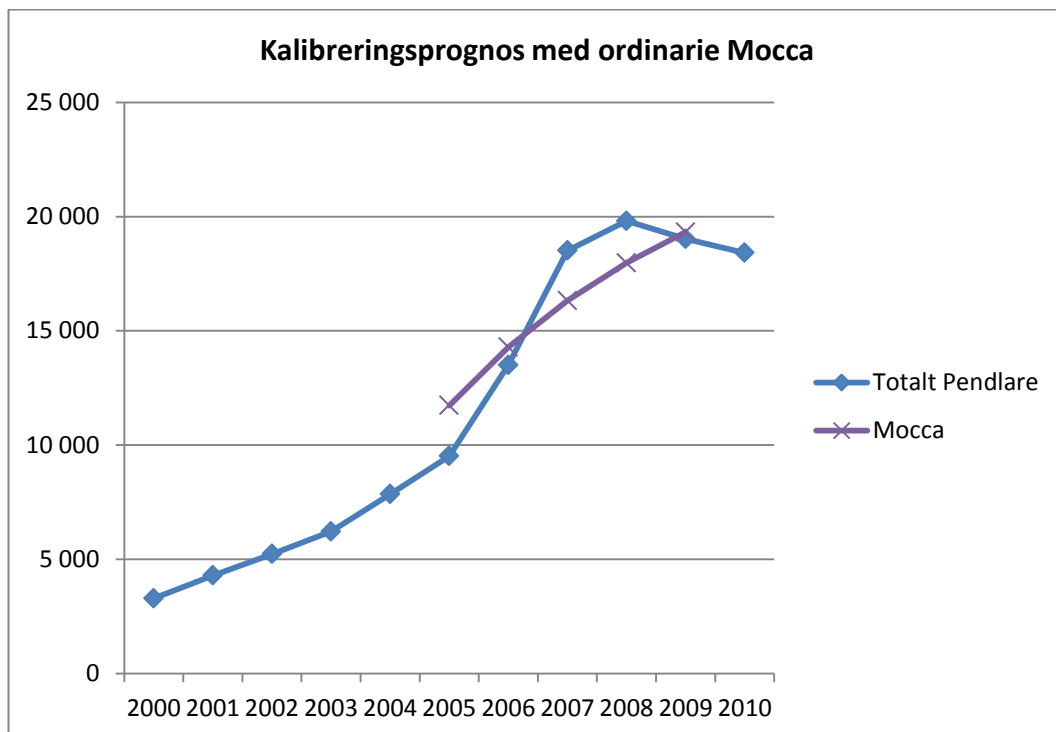
vånarna i regionen att byta arbets- och bostadslokalisering så att de kan utnyttja tillgänglighetsförbättringen från öppnandet av bron.

Ökningen av bostadsprisskillnaderna mellan 2001 och 2006 innebar att de danska bostäderna successivt ökade mer i värde än de på den svenska sidan i regionen. Detta innebar i sin tur att en person boende på den danska sidan kunde tjäna betydande summor ekonomiskt i form av lägre boendekostnader på att flytta till den svenska sidan. Vidare, eftersom löneskillnaderna mellan danska och svenska sidan är sådan att lönen är högre på den danska sidan är det också fördelaktigt för de personer som har flyttat till den svenska sidan att arbeta kvar på den danska sidan. Detta i sin tur resulterar i arbetspendling över Öresund.

Eftersom både löner och bostadspriser är högre på den danska sidan så är det arbetspendling över sundet av personer boende på den svenska sidan som helt dominerar arbetspendlingen. Från 2006 och framåt har arbetspendlarna till minst 95 procent bestått av personer som har haft sin bostad på den svenska sidan.

Validering

Den ordinarie versionen av Mocca är kalibrerad så att den ska stämma bra överens med antalet pendlare över sundet år 2009. Vid kalibreringen användes löne- och bostadsprisskillnader som låg nära de faktiska nivåerna för 2006, löneskillnad = 32,5 tkr och bostadsprisskillnad = 1310 tkr. Figur 2 nedan visar resultatet av Moccas kalibreringskörning.

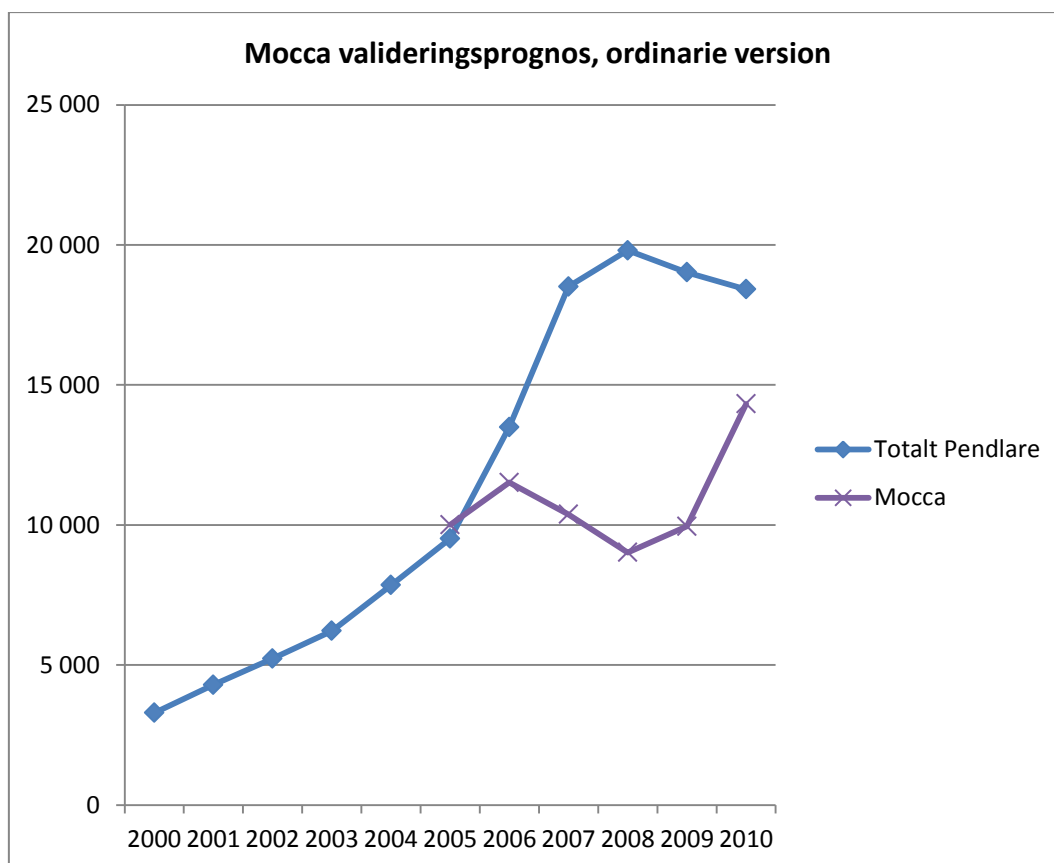


Figur 2: Prognos hämtad från kalibreringen av ordinarie versionen av Mocca.

Att kalibreringen av Mocca är gjord mot antalet pendlare 2009 samtidigt som skillnaderna i löner och bostadspriser som användes var högre än de faktiska skillnaderna för åren 2007-2010 (med undantag för löneskillnaderna år 2010), innebär att antalet pendlare antalet pendlare med automatik kommer att ligga under de faktiska nivåerna i valideringsprognosen där de faktiska löne- och bostadsprisskillnaderna används.

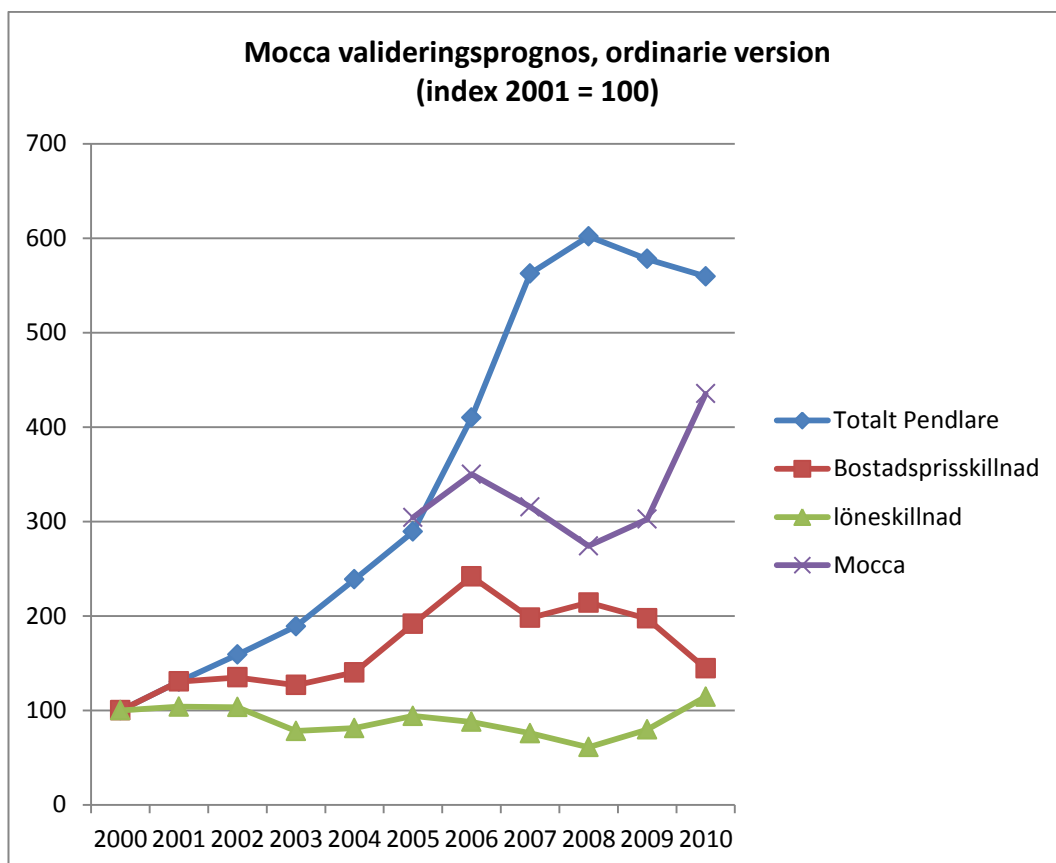
Valideringsresultatet beror därmed i relativt låg grad på hur nära Mocca prognostiserar de faktiska pendlingsnivåerna. Det viktiga är hur Mocca reagerar på förändrade skillnader i löner och bostadspriser. Om Mocca reagerar på ett lämpligt sätt på förändrade skillnader så räcker det med en omkalibrering av modellen för att ge prognoser som ligger nära de faktiska pendlingsnivåerna under valideringsperioden.

Figur 3 nedan visar Moccas prognos för valideringsperioden 2006-2010 när de faktiska skillnaderna i löner och bostadspriser används.



Figur 3: Valideringsprognos, ordinarie version av Mocca.

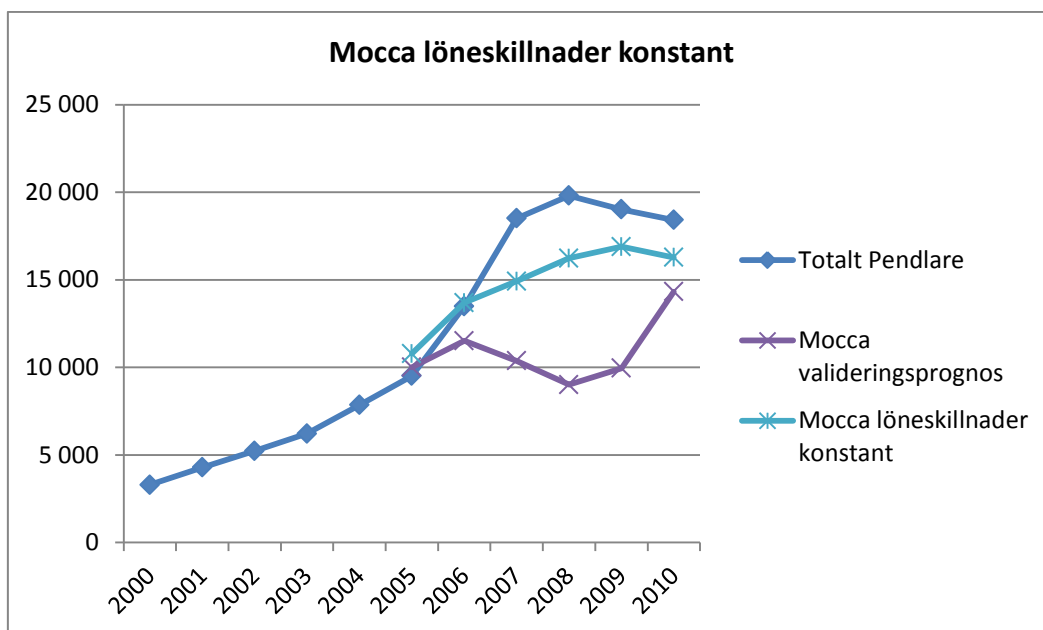
Det är uppenbart att Moccas pendlingsprognos avviker i princip direkt när valideringsperioden startar. För att få en bättre bild av vad som sker har utvecklingen av löne- och bostadsprisskillnader lagts in i nästa diagram (som är ett indexdiagram, med 2001 som basår).



Figur 4: Indexdiagram motsvarande figur 3 med serierna för löne- och bostadsprisskillnader inlagda.

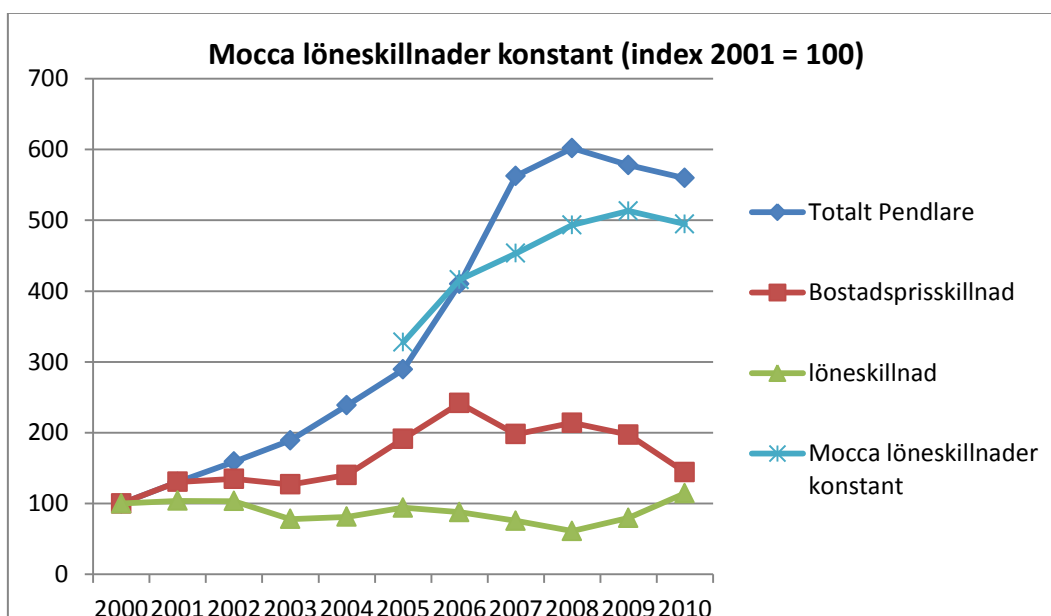
Jämfört med kalibreringsprognosen i figur 2 ovan så ligger Moccas prognos nu under den faktiska pendlingsnivån för 2006. Bostadsprisskillnaderna ligger i denna prognos något över skillnaderna i kalibreringsprognosen, det kan inte förklara att antalet pendlare i Moccas valideringsprognos ligger under värdet i kalibreringsprognosen. Löneskillnaden i kalibreringsprognosen för 2006 var 32,5 tkr och 30,2 tkr i valideringsprognosen. Det är en förändring som kan ge en lägre nivå på pendlingen i valideringsprognosen jämfört med kalibreringsprognosen. Vidare så följer den fortsatta pendlingsutvecklingen i valideringsprognosen i stora drag utvecklingen av löneskillnaderna. Mellan åren 2006-2008 minskar löneskillnaderna samtidigt som valideringsprognosen visar på sjunkande pendling. Mellan åren 2008-2010 ökar löneskillnaderna och pendlingen i valideringsprognosen ökar.

Valideringsprognosen antyder att Moccas känslighet för löneskillnader är för stor. För att testa den hypotesen kördes två varianter av prognoser fram som var kombinationer av kalibrerings- och valideringsprognoserna. I den första varianten användes samma löneskillnadsutveckling som i kalibreringsprognosen (konstant 32,5 tkr) men utvecklingen av bostadsprisskillnaderna hämtades från valideringsprognosen. I den andra varianten gjordes tvärtom, utvecklingen av bostadsprisskillnaderna hämtades från kalibreringsprognosen (konstant 1310 tkr) medan löneutvecklingen togs från valideringsprognosen. Figur 5 nedan visar den först varianten.



Figur 5: variant valideringsprognos Mocca med konstanta löneskillnader (32,5 tkr) från kalibreringsprognosen.

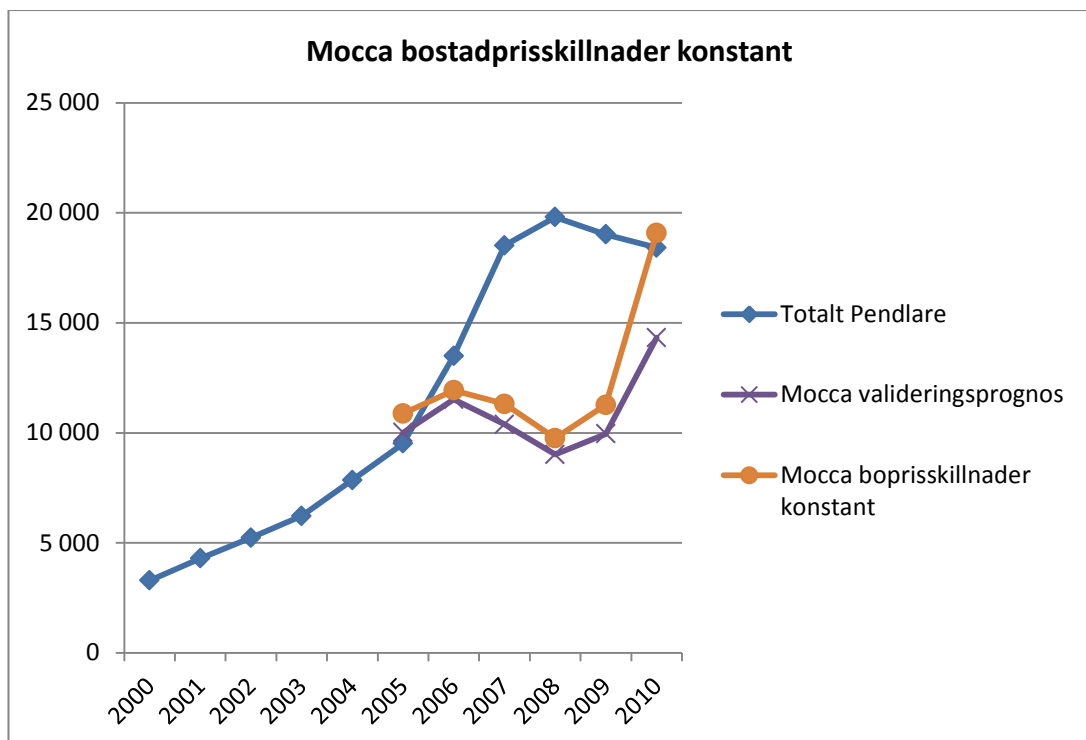
I detta fall ger Mocca en prognos som ligger betydligt närmare kalibreringsprognosen, men också det faktiska utfallet för arbetspendlingen över sundet. Figur 6 nedan visar motsvarande indexdiagram med utvecklingen av löne- och bostadsskillnaderna inlagda.



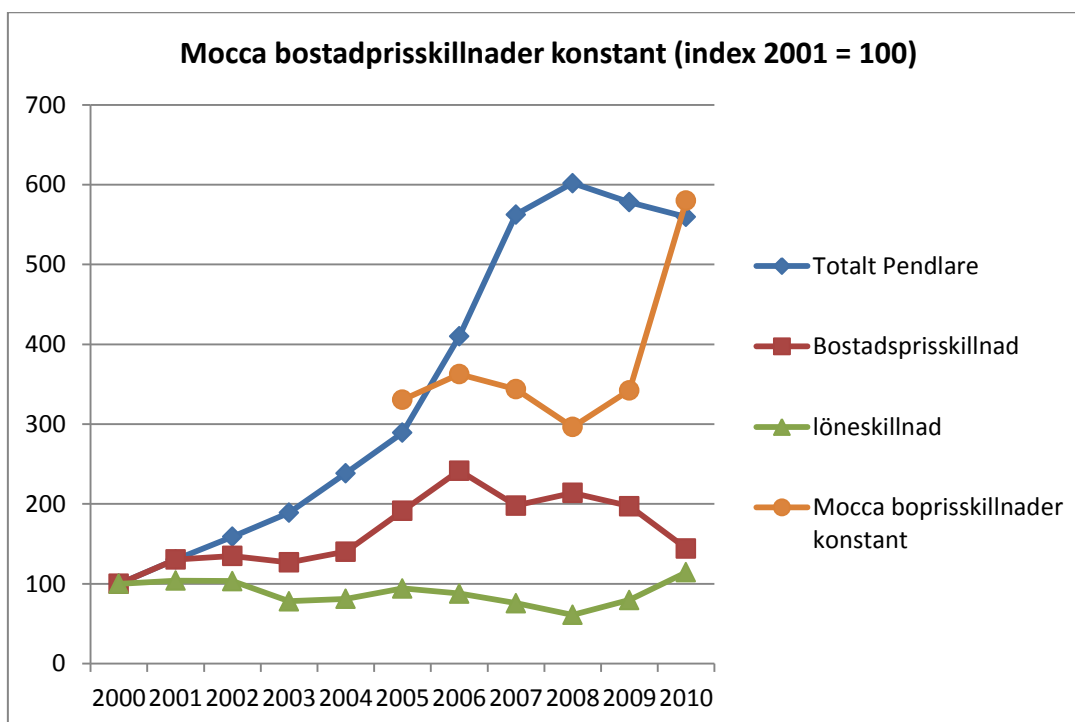
Figur 6: Indexdiagram motsvarande figur 5 med serierna för löne- och bostadsprisskillnader inlagda.

Figur 7 (och figur 8, med motsvarande indexdiagram) nedan visar den variant på valideringsprognos där bostadprisskillnaderna istället hålls konstanta på den nivå som användes i kalibreringsprognosen (1310 tkr).

Det är uppenbart från figur 7 att när bostadprisskillnaden hålls konstant på nivån från kalibreringsprognosen, så följer prognosen pendlingsutvecklingen från valideringsprognosen mycket nära. Enbart för det sista året 2010 avviker denna variant markant från valideringsprognosen. Slutsatsen är att Mocca i sin nuvarande version visar en för stor känslighet för förändrade löneskillnader mellan danska och svenska sidan av regionen.



Figur 7: variant valideringsprognos Mocca med konstanta bostadprisskillnader (1310 tkr) från kalibreringsprognosen.



Figur 8: Indexdiagram motsvarande figur 7 med serierna för löne- och bostadprisskillnader inlagda.

Omestimering av Mocca

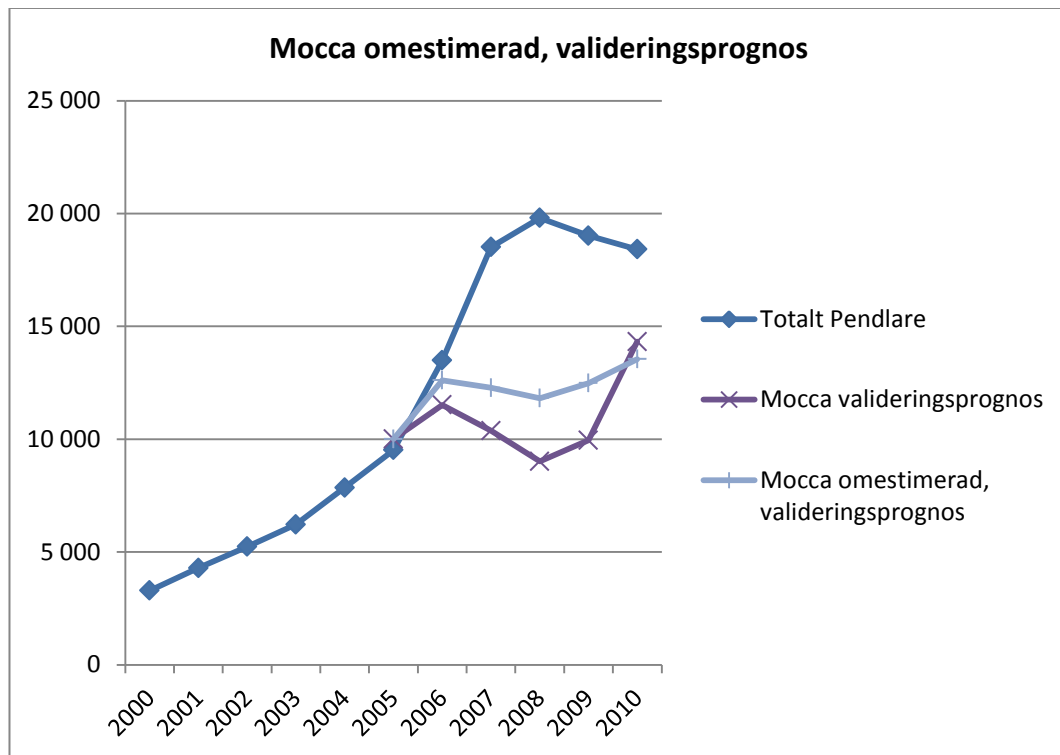
Som nämndes i inledningen så utnyttjades inte hela datamaterialet vid den ursprungliga estimeringen av Mocca. Det lämnar utrymme för hypotesen att den ovan konstaterade för höga känsligheten för löneskillnader kan bero på detta och att en omestimering skulle ge bättre känsligheter.

Tabell 2 nedan visar förändringen av känsligheten mätt i förändring av parametervärden för både löne- och bostadprisskillnader efter omestimering på det fullständiga datamaterialet.

Tabell 2: Förändring av parametervärden vid omestimering av Mocca på den fullständiga datamängden för perioden 2001-2006.

Parameter	Förändring (%)
Bopris, Svenskar	-14
Bopris, Danskar	57
Lön, Svenskar	-65
Lön, Danskar	-33

Tabellen visar en betydande minskning av känsligheten för löneskillnader för danskar men framförallt personer av svensk nationalitet, men också en betydande förändring av känsligheterna för bostadssprisskillnader. Figur 9 nedan visar hur valideringsprognosen ändras när de nya parametervärdena används i modellen.



Figur 7: valideringsprognos med omestimerad Mocca.

Den omestimerade modellen visar på en förbättring när det gäller känsligheten för löneskillnader. Fortfarande finns dock indikationer på att lönekänsligheten är för stor. Till exempel så hamnar fortfarande pendlingen 2006, med den omestimerade modellen, markant under pendlingsnivån som ges av kalibreringsprognosen (för den gamla modellen), och pendlingsutvecklingen vänder 2008 när löneutvecklingen vänder. Så slutsatsen lutar åt fortsatt för hög känslighet för löneskillnader även efter omestimering.

Övrigt arbete

Utöver att validera mot antalet arbetspendlare totalt över sundet så har kontroll utförts av hur väl modellens fördelning av pendlare mellan svenska och danska boendesidorna i regionen stämmer mot det verkliga utfallet. I valideringsprognosen för 2010 bodde 92,2 procent av pendlarna på den svenska sidan. I det verkliga utfallet för 2010 var andelen 96,1 procent. Mocca underskattar alltså andelen pendlare som bor på den svenska sidan något. Dock så behöver denna avvikelse inte vara speciellt signifikant eftersom nivån på pendlingen i valideringsprognosen låg markant under det verkliga utfallet. För att pendlingen ska minska i Mocca måste skillnaderna mellan de två sidorna av sundet relativt sett

minska Det implicerar i sin tur att andelen pendlare som bor på svenska sidan ska vara lägre för pendlingsnivån som gäller för valideringsprognosen, jämfört med den verkliga pendlingsnivån.

En annan jämförelse som är intressant är fördelningen av pendlare på nationalitet (medborgarskap). Mocca är segmenterad i separata modeller för danskar och svenskar. Den jämförelsen är dock svår att göra eftersom inga lättillgängliga statistiskuppgifter finns över medborgarskap. Däremot finns det statistiskuppgifter om pendlarnas fördelning över födelseland. Ett problem är att födelseland och nationalitet i form av medborgarskap inte är riktigt samma sak. Ett annat problem är att 23 procent av pendlarna är födda utanför Danmark eller Sverige medan en individ i Mocca är antingen är dansk eller svensk. Vid estimeringen av Mocca löstes det genom att en person som inte hade danskt eller svenskt medborgarskap tilldelades nationalitet efter dennes boendesida av sundet vid det tillfälle som person först dök upp i data. Något liknande går dock inte att göra med de uppgifter som finns om den verkliga pendlingens fördelning efter födelseland. Detta är betydande svårigheter, men de undersökningar som utfördes i samband med estimeringen av Mocca visade att (1) födelseland eller medborgarskap som grund för en nationalitetsuppdelning inte har avgörande betydelse, och (2) exakt hur de personer hanteras som antingen är födda utanför Danmark eller Sverige, eller inte har danskt eller svenskt medborgarskap, inte har en avgörande betydelse för samband relaterade till arbetspendlingen över sundet. Därför görs, trots allt, en jämförelse mellan pendlarnas fördelning över "nationalitet" mellan det verkliga utfallet och Moccas valideringsprognos. Tabellen nedan visar resultatet. För det verkliga utfallet har personer födda utanför Danmark eller Sverige först uteslutits innan fördelningen över födelseland har beräknats. För prognosen har Moccas nationalitetsbegrepp använts rakt av.

	Verkligt utfall	Valideringsprognos
Svensk	52%	44%
Dansk	48%	56%
Totalt	100%	100%

Det kan synas som att skillnaderna mellan verkligt utfall och valideringsprognos enbart handlar om mindre rörelser runt nivån 50 procent, men de innebär att det går 0,94 danska pendlare per svensk pendlare i det verkliga utfallet för år 2010, medan Mocca prognostiserar 1,32 danska pendlare per varje svensk pendlare. Det är en skillnad som är svår att förklara med enbart definitionsproblemen beskrivna ovan. Skillnaden mellan utfall och prognos är för nationalitetsfördelningen också svårare att koppla till skillnader i totalnivån för pendlingen över sundet, än vad som var fallet för fördelningen över pendlarnas boendeland.

Både en eventuell snedfördelning av pendlarnas boendeland och den märkbara snedfördelningen av pendlarnas nationalitet kan relativt enkelt justeras vid en omkalibrering. Den

sistnämnda snedfördelningen kan dock ha samma orsak som problemet med den höga känsligheten för lönekostnader. Detta diskuteras mer i nästa avsnitt.

Mocca kör fram en prognos genom att modellen tillämpas inkrementellt (pivot point), genom att en observerad övergångsmatrix används som bas för prognosen. Övergångsmatrisen ger sannolikheten att en person som bor och arbetar i en viss kombination av områden innevarande år, kommer att bo och arbeta i en viss (eventuellt annan) kombination av bostads- och arbetsområden följande år. Den observerade övergångsmatrisen justeras i nästa steg utifrån de värden på prognosvariablerna som används i prognosen. I nuvarande version av Mocca används den observerade övergångsmatrisen mellan åren 2005 och 2006. Den innehåller dock många par av kombinationer av bostads- och arbetsområden som helt saknar byten mellan de två åren. För att utesluta hypotesen att detta faktum har påverkat den faktiska känsligheten för löneutveckling i modellen (vid prognoser) har en mer fullständig övergångsmatrix testats som är summan av alla fem övergångsmatriser i estimationsdata för perioden 2001-2006. Den övergångsmatrisen har färre tomma celler. En prognos med den övergångsmatrisen gav endast marginella förändringar för valideringsprognosen, jämfört med den övergångsmatrix som används i nuvarande version av modellen. Hypotesen att eventuella brister i den observerade övergångsmatrisen ligger bakom de höga känsligheterna för löneutveckling förefaller därmed mindre sannolik.

Ett alternativt sätt att testa hur väl den observerade övergångsmatrisen fungerar i modellen är att jämföra de tidigare valideringsprognoserna med en syntetisk tillämpning av modellen. I en syntetisk tillämpning modellberäknas övergångsmatrisen fullständigt och inga observerade estimeringsdata används i modellen. En test av en sådan syntetisk körning av valideringsprognosen gav enbart marginella avvikelser i antal arbetspendlare jämfört med när observerade övergångsmatriser användes (både jämfört med 2005/2006 och den summerade matrisen för åren 2001-2006). Detta stärker slutsatsen från föregående stycke.

Det har förekommit diskussioner om möjligheten att använda Mocca för prognoser i andra gränsregioner. Att Mocca tillämpas inkrementellt är då en nackdel. En ny observerad matris bör i så fall tas fram för den aktuella gränsregionen. Detta kräver i princip mikrodata på individnivå från båda sidor av gränsen. Detta har varit möjligt i Öresundsregionen på grund av det särskilda samarbete som finns mellan SCB och Danska statistikbyrån i regionen. I andra regioner finns dock inte sådana samarbeten i dagsläget. Om Mocca kan tillämpas syntetiskt bortfaller dock behovet av mikrodata på denna nivå. Mocca skulle då istället kunna flyttas genom att kalibrera in modellen till den nya gränsregionen med hjälp av tillgängliga data på aggregerad nivå. På det sättet var testet med syntetisk tillämpning av Mocca lovande. Dock krävs mer tester innan det slutligen går att avgöra om Mocca fungerar bra i en syntetisk tillämpning.

Kalibreringsrutinen i Mocca har haft vissa begränsningar. Bland annat förutsätts ett prognosscenario där löne- och bostadspri skillnader hålls konstanta för alla år som ingår i kalibreringsprognosen. Kalibreringsmålen som kalibreringen syftar till att återskapa har enbart kunnat sättas för ett specifikt år (det sista året som ingår i prognosscenariet). Dessa begränsningar har åtgärdats vilket öppnar för mer flexibla kalibreringar där löne- och bo-

stadsprisskillnader kan variera år för år och kalibreringsmålen kan gälla separat för ett godtyckligt antal år i scenariet.

Orsaker till den höga känsligheten för löneskillnader

Ovan kunde konstateras att känsligheten för löneskillnader var hög. Resultatet av testerna med en annan variant av övergångsmatris och syntetisk tillämpning av modellen innebär att problemet troligen inte beror på en felaktig implementering av den estimerade modellen. Därmed måste orsaken sökas vid estimering av modellen eller dess specifikation. Omestimering med den fullständiga datamängden för estimeringsperioden 2001-2006 innebär visserligen att känsligheten mildrades, men problemet försvann inte. Totalt sett betyder detta att det är troligt att modellens specifikation avseende arbetsplatsbyten behöver förbättras.

Konkret betyder detta att genomsnittlig löneskillnad mellan den danska och svenska sidan av Öresundsregionen troligen inte är den enda variabel som har betydelse för byten av arbetsplats. I så fall är det troligt att de estimerade parametererna för löneskillnadsvariabeln innehåller effekter från andra betydelsefulla variabler för arbetsplatsbyten som är korrelerade med löneskillnadsvariabeln men inte ingår i modellen. Detta är en situation som typiskt ger försämrade prediktionsförmåga. Det räcker med att någon av dessa variabler, som inte är med i modellen, ändrar sitt värde under valideringsperioden (jämfört med estimeringsperioden), för att modellen ska ge en otillfredsställande prediktion.

Ett annat sätt där modellens specifikation kan ge problem är att skilda grupper av individer reagerar olika på variabler som ingår i modellen eller viktiga variabler som inte ingår i modellen. Detta betyder att modellen bör ytterligare segmenteras i delmodeller. En hypotes som täcker både saknade variabler och ett ökat segmenteringsbehov är att unga personers agerande på arbetsmarknaden mer handlar om att de är beredda att flytta eller ta ett arbete i ett mindre tillgängligt område för att få ett arbete. Så som Mocca i nuläget är specificerad beskriver modellen troligare äldre mer etablerade personers agerande på arbetsmarknaden. För dessa personer är möjligheten att byta arbetsplats för att höja lönen troligen viktigare än att undkomma arbetslöshet. Löneskillnadsvariabeln är därmed mer relevant för dessa personer.

Om vi ser på det faktiska utfallet för unga och äldres arbetspendling över Öresund under perioden 2001-2011 så kan vi dra följande slutsatser. Åldersgruppen 16-24 år hade sin topp mätt i andel av pendlarna, 15-17 procent, under åren 2006-2008 för att sedan sjunka relativt drastiskt ned till 9 procent av pendlarna år 2011. Det är samma andel som den åldersgruppen hade år 2001. För äldre pendlare i åldern 35-44 år gällde att de för hela perioden 2001-2008 låg tämligen konstant på en andel av cirka 25 procent. För de senaste åren 2009-2011 har de ökat sin andel betydligt till cirka 30 procent. Totalt sett visar det här på ett tämligen intrikat mönster för den unga gruppen och för den något äldre men mer etablerade åldersgruppen. Finanskrisen 2008/2009 ser ut att vara en viktig händelse för dynamiken i båda grupperna under perioden 2001-2011. Om det är så att drivkrafterna för dessa två grupper arbetssökande ser olika ut (undkomma arbetslöshet respektive op-

timera lönen genom arbetsbyten) är det troligt att det har resulterat i en snedvriden estimering av lönekänsligheten under estimeringsperioden 2001-2005, och en felaktig påverkan av löneskillnadsvariabeln på prognosen för pendlingen under valideringsperioden 2006-2010.

Den sammantagna slutsatsen är att om modellen ska användas för att under varierande förhållanden på arbetsmarknaden prognostisera arbetspendlingen över Öresund bör dess specifikation utvecklas så att den mer adekvat beskriver förhållanden på arbetsmarknaden.

Med detta sagt så är det också av intresse att diskutera hur modellen bör användas i sin nuvarande version med hänsyn till valideringsresultaten. I princip samtliga prognoser där Mokka har använts fram till och med idag, så har samma värden på löne- och bostadsprisskillnader använts som användes för kalibreringen, det vill säga löneskillnad 32,5 tkr och bostadsprisskillnad 1310 tkr. Det innebär i princip att de variablerna har hanterats som ett par extra kalibreringskonstanter. En motivering till detta är prognosernas syfte har varit att analysera tillgänglighetsförändringar. Både löne- och bostadsprisskillnad har det gemensamt att de är svåra variabler att prognostisera så att de har hållits konstanta på de ovan givna värdena kan också ses som att de är icke-prognostiserade variabler som har ansatts typvärden för att i praktiken utesluta dem från prognoserna. Den höga konstaterade känsligheten för löneskillnader är ett argument för att åtminstone löneskillnaden också i fortsättningen bör hanteras på detta sätt. Det använda värdet på 32,5 tkr ligger i övre delen av de observerade värdena för perioden 2001-2011, men det är också svårt att hävda att det finns någon entydig trend för utveckling av lönerna under perioden – till exempel konvergens mot lägre löneskillnader i regionen. Att i rena trafikanalyser fortsätta att konstanthålla löneskillnaden i prognoser förefaller som ett relativt rimlig förhållningssätt. Dock bör det noteras att detta utesluter att studera det dynamiska förloppet för pendlingsutvecklingen på kort sikt.

WSP och GENIVAR har gått samman och bildar tillsammans ett av världens ledande analys- och teknikkonsultföretag. Vi erbjuder tjänster för hållbar samhällsutveckling inom Hus & Industri, Transport & infrastruktur och Miljö & Energi. Bredd och mångfald kännetecknar våra medarbetare, kompetensområden, kunder och typer av uppdrag. Tillsammans har vi 15 000 medarbetare på över 300 kontor i 35 länder. I Sverige har vi omkring 2 500 medarbetare.

Vår verksamhet bedrivs inom WSP Analys & Strategi, WSP Brand & Risk, WSP Byggprojektering, WSP Environmental, WSP International, WSP Management, WSP Process, WSP Samhällsbyggnad och WSP Systems.

Bredd och mångfald kännetecknar våra medarbetare, kompetensområden, kunder och typer av uppdrag. Vi är *United by our difference*.