

Metodevedlegg til kapittel 10¹

Dette vedlegget omtaler utvalgte metoder, antagelser og feilkilder for noen fagartikler som er omtalt i avsnittet «Makroøkonomiske effekter av det indre marked». Modellene som er omtalt her er gravitasjonsmodeller (gravity-modeller), den syntetiske kontrollmetoden og forskjell-i-forskjell-modeller (difference-in-difference-modeller).

Som nevnt i kapittelet er det vanskelig å måle effekten av deltagelsen i det indre marked på økonomiske variable som BNP, handel mv., selv om man tar i bruk sofistikerte økonometriske metoder. Det skyldes blant annet at det er nødvendig å gjøre antagelser om det kontrafaktiske forløpet. Alle slike anslag er derfor forbundet med usikkerhet, noe det er grunn til å understreke.

Det er likevel verdt å merke seg at de ulike modelltypene, som er bygget opp på nokså forskjellige måter, jevnt over gir anslag for effekten av det indre marked i samme størrelsesorden. Alle artiklene som er omtalt i kapittel 10 finner for eksempel positive effekter på handel og BNP. Når det gjelder BNP-effekten, finner artiklene i gjennomsnitt at det indre marked har økt medlemslandenes BNP med rundt 3 til 6 prosent. Det betyr altså at verdiskapingen hvert år er 3 til 6 prosent høyere enn den ellers ville vært. Som regel innebærer høyere verdiskaping høyere materiell velferd for en befolkning. Men å oversette økt handel og BNP til velferdsgevinster krever vurdering av blant annet fordelings effekter, og vi minner om at BNP ikke er et velferdsmål i seg selv.

Gravitasjonsmodeller

Gravitasjonsmodeller tar utgangspunkt i at internasjonal handel mellom to land særlig er påvirket av landenes størrelse og avstanden mellom dem. Det blir dermed en parallell til gravitasjonsligninger i fysikken, uten at parallellen skal gjøres for mye ut av. Modellene bygger på standard økonomisk teori. Det finnes også mye empirisk støtte for at størrelse og avstand, eller transaksjonskostnader mer generelt, har betydning for handel.²

I utgangspunktet er dette en beskrivelse av en mer eller mindre stabil sammenheng i internasjonal økonomi. Men slike modeller har siden Tinbergen (1962) også blitt brukt til å analysere bilaterale handelsstrømmer. Man kan kombinere estimater fra gravitasjonsmodeller med handelspolitiske eksperimenter for å anslå implisitte velferdsendringer av politiske tiltak, som det indre marked er et eksempel på. Hovedpoenget er altså å komme fram til en koeffisient som anslår hvor stor effekt politikken har hatt, og på den måten blir produksjon og geografiske forhold kontrollvariabler.

Metoden er brukt blant annet i Mayer m. fl. (2019).

¹ Utvalget takker Edda Torsdatter Solbakken, forsker i Statistisk sentralbyrå, som har lest gjennom tidligere utkast til dette vedlegget og gitt nyttige innspill til framstillingen.

² Men det finnes også modeller der økonomiens størrelse i utgangspunktet er underordnet. I et Hecksher-Ohlin-rammeverk er det særlig forskjell på faktortilgang (arbeid, kapital, mv.) mellom land som bestemmer handelsstrømmer.

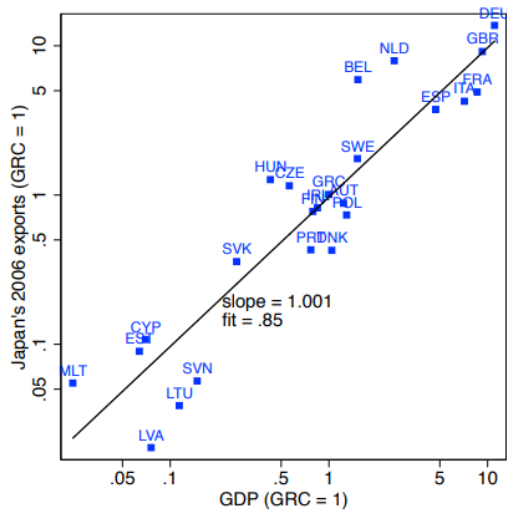
Figurene under er hentet fra Head og Mayer (2013) og illustrerer de to sentrale sammenhengene for Japan og Frankrike. Handel øker med BNP-størrelse og avtar med avstand/transaksjonskostnader.

I praksis relaterer gravitasjonsmodeller logaritmen av pengeverdien av handel mellom to land til logaritmen av deres respektive BNP, dessuten en kompositt som måler barrierer og insentiver til handel mellom landene, og vilkår som måler barrierer for handel mellom hver av dem og resten av verden. Estimeringsmetoden er som regel minste kvadraters metode (OLS).

To begrensninger knyttet til bruk av gravitasjonsmodeller for å estimere virkningen av en handelsavtale bør nevnes. For det første kan handelsavtaler være endogene variabler. Det vil si at årsakssammenhengen mellom inngåelsen av en handelsavtale og handelsstrømmene kan gå fra sistnevnte til førstnevnte. I så fall bestemmes handelsavtalen av, i stedet for å bestemme, handelsstrømmene. Dette kan påvirke estimatene, og omfanget av skjevheten kan være ganske stort. Videre er det mange handelsavtaler o.l. som er opprettet med andre formål enn bare økt handel, og hvor «ikke-tradisjonelle» gevinster kan være en del av motivasjonen for å inngå avtalen. Slik er vanskelig å fange opp i et gravitasjonsmodell-rammeverk, men den kritikken gjelder også for andre økonometriske metoder.

Figure 1 – Trade is proportional to size

(a) Japan's exports to EU, 2006



(b) Japan's imports from EU, 2006

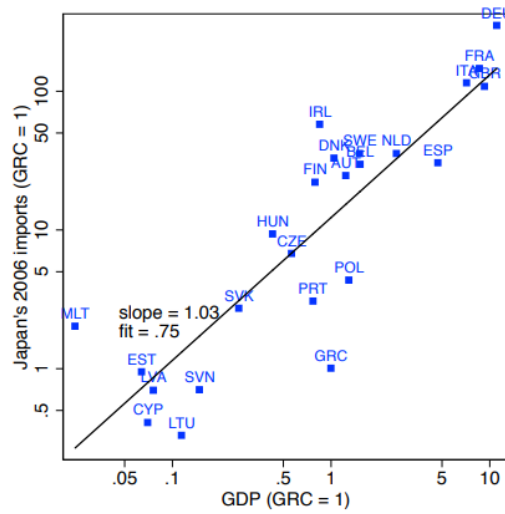
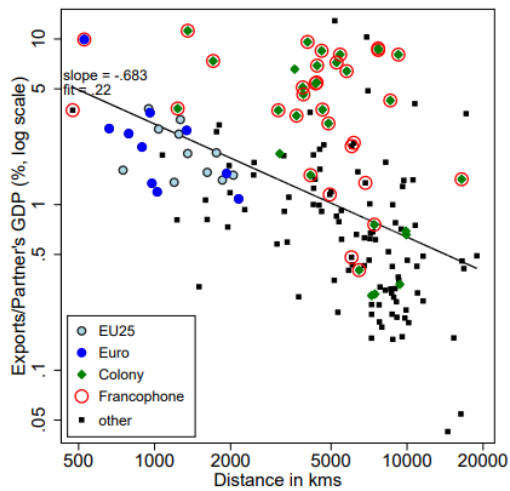
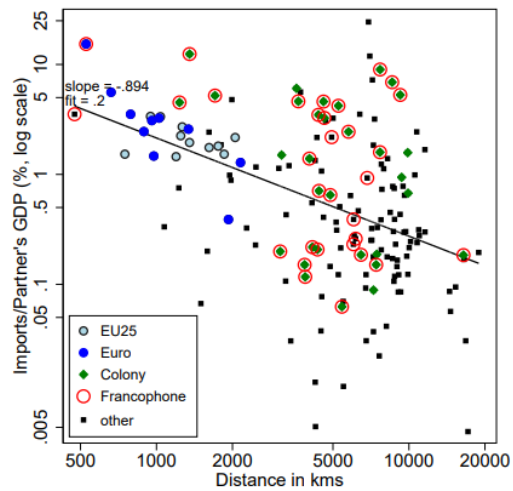


Figure 2 – Trade is inversely proportional to distance

(a) France's exports (2006)



(b) France's imports (2006)



Den syntetiske kontrollmetoden

Lehtimäki og Sondermann (2022) analyserer hvordan det indre markedet har påvirket EU-landene som dannet det indre marked tidlig i 1993, dvs. Belgia, Danmark, Frankrike, Tyskland, Hellas, Irland, Italia, Luxembourg, Nederland, Portugal, Spania og UK ved å bruke

den *syntetiske kontrollmetoden*. Metoden ble utviklet tidlig på 2000-tallet og er de siste 20 årene blitt mye brukt til policy-analyser.³

Campos m. fl. (2014) bruker også den syntetiske kontrollmetoden i sitt oppsett.

Den syntetiske kontrollmetoden er en videreutvikling av forskjeller-i-forskjeller-metoden⁴. Grunntanken bak forskjeller-i-forskjeller-metoden er å sammenligne endringen over tid for enheten eller gruppen som blir utsatt for intervensjonen eller «behandlingen» vi er interessert i – den behandlede enheten – med endringen over tid i en sammenlignbar enhet eller gruppe som ikke har blitt utsatt for behandlingen – kontrollgruppen.

For at vi kan konkludere med at metoden gir oss effekten av behandlingen må man ha tiltro til den underliggende antagelsen om at utviklingen i de to gruppene ville vært lik hvis ikke den behandlede enheten hadde blitt utsatt for behandlingen. Dette er det selvfølgelig ikke mulig å bekrefte med sikkerhet, siden den behandlede enheten ble behandlet. I stedet er det standard å teste for om utviklingen var lik *før* den behandlede enheten ble behandlet, og deretter argumentere for at det ville fortsatt slik om behandlingen ikke hadde funnet sted.

I mange tilfeller er det vanskelig å finne en gruppe eller enhet som utvikler seg som den behandlede enheten i fravær av behandlingen. I tilfelle vårt må vi finne land som utviklet seg som Norge ville gjort om Norge ikke ble med i det indre markedet. Det er vanskelig å se for seg et land som er så likt Norge at vi tror på antagelsen om at Norge hadde utviklet seg akkurat som dette landet hvis Norge ikke hadde meldt seg inn i indre marked.

Den syntetiske kontrollmetoden utvider muligheten for kontrollgruppen. I stedet for å finne et land som er så sammenlignbart at man tror på antagelsen om at utviklingen i landet representerer utvikling Norge ville hatt om Norge ikke meldte seg inn i det indre marked, kan man konstruere et syntetisk-kontroll-land ved å ta et vektet gjennomsnitt av faktiske land. Vekten ulike land får i syntetisk-kontroll-landet settes slik at utviklingen i Norge og syntetisk-kontroll-landet blir likest mulig i perioden før Norge melder seg inn i det indre markedet. Som i forskjeller-i-forskjeller-metoden kan man ikke teste empirisk om utviklingen i Norge etter Norge meldte seg inn i det indre marked ville vært lik utviklingen i syntetisk-kontroll-landet. I stedet for må man argumentere for at det ikke har skjedd noe i Norge eller landene som utgjør syntetisk-kontroll-landet, annet enn at Norge har meldt seg inn i det indre markedet, som kan ha ført til at utviklingene ikke ville vært lik.

Metoden forutsetter at man har data om et sett med land/enheter. Settet kaller vi $J + 1$. Enheten(e) som skal undersøkes ($j=1$), Norge i vårt tilfelle, går gjennom en «behandling». De resterende landene, $j=2, \dots, J+1$ går ikke gjennom behandlingen.

Det kontrafaktiske «ikke-behandlede» forløpet beregnes ved hjelp av land som ikke ble utsatt for behandlingen, heretter kalt donorutvalget. Disse settes sammen på en spesiell måte og danner det som altså kalles en syntetisk kontrollgruppe. Abadie m. fl. (2015) definerer den syntetiske kontrollgruppen som et vektet gjennomsnitt av landene i donorutvalget. Det nyskapende med metoden er hvordan vektene utarbeides.

³ Viktige bidragsytere var Alberto Abadie og Javier Gardeazabal ved Harvard University. Metoden som er referert til i dette avsnittet følger deres opprinnelige oppsett i en artikkel fra 2003, beskrevet her: <https://pubs.aeaweb.org/doi/pdfplus/10.1257/jel.20191450>

⁴ Engelsk: Differences-in-differences. Ofte omtalt som diff-in-diff-metoden, selv på norsk.

Metoden gir en $(J \times 1)$ -vektor med vektor $W = (w_2, \dots, w_{J+1})'$. Vektene er mellom 0 og 1 og summerer seg til 1. Verdien på vektene for hvert enkelt land velges slik at *egenskapene til den behandlede enheten skal ligne mest mulig på egenskapene til den syntetiske kontrollgruppen*.

X_1 er en $(k \times 1)$ -vektor som inneholder verdiene på egenskapene før behandling for den behandlede enheten. Det er disse egenskapene man ønsker å matche så godt som mulig når man skaper en kunstig kontrollgruppe. Som i medisinske forsøk er det ønskelig at «pasientene» ligner så mye som mulig før behandling når man skal sjekke effekten av en medisin kontra ingen behandling.

Tilsvarende er X_0 en $(k \times J)$ -matrise med verdier for de samme variablene for alle enhetene i donorutvalget. *Metoden velger en syntetisk kontroll, W^* (altså et sett med vektor), som minimerer avvikene mellom den behandlede enheten og den syntetiske kontrollgruppen, dvs. at man minimerer vektoren $X_1 - X_0W$.*⁵

Forskerne må gjøre tre valg som vil bestemme hvilke vektorer ulike land får, som igjen vil påvirke resultatene:

- 1) hvilken tidsperiode som defineres som før-behandlings-perioden,
- 2) hvilke egenskaper man ønsker at skal være like mellom den syntetiske kontrollen og Norge i før-behandlings-perioden
- 3) hvilke land man inkluderer i donorutvalget.

Tidsrommet som dekkes er 1964 til 2014. Perioden 1964-1992 defineres som før-behandlings-perioden, perioden 1993-2014 defineres som behandlingsperioden.

Hvilke egenskaper man ønsker at skal være like mellom den syntetiske kontrollen og Norge i før-behandlings-perioden, avhenger av skjønn og økonomisk teori. Lehtimäki og Sondermann (2022) har først brukt variablene fra en standard nyklassisk vekstmodell: Kapitalstokk, befolkning og mål på humankapital for å måle arbeid og kapital. Videre investeringsrater, totalfaktorproduktivitet og BNP per innbygger. Deretter inkluderes andre faktorer som også kan påvirke økonomisk vekst, som styrke på institusjoner, åpenhet for handel mv.

Av flere grunner er det ønskelig at donorutvalget inneholder land som antas å ligge noenlunde nært det behandlede landet for en del av disse variablene. Rent matematisk kan man se for seg at snittet av et lutfattig land og et styrtrikt land kunne brukes som kontrollgruppe for et mellominntektsland, men dette gir fort merkelige resultater.

Forfatterne bruker land som ikke har blitt medlem av det indre marked etter 1993 som donorutvalg til kontrollgruppe. Å bruke land som senere blir medlem av det indre marked ville svekket kredibiliteten til antagelsen om at den syntetiske kontrollen utvikler seg som Norge ville gjort om Norge ikke var medlem av det indre markedet. Utvalget med kontrolland er begrenset til lignende utviklede land, dvs. alle land som var medlemmer av OECD ved tidspunktet for opprettelsen av det indre marked, og ikke var blant medlemslandene for det indre marked (direkte eller per tredjelandsavtale). Dette er Australia, Canada, Israel, Japan, New Zealand og USA. Forfatterne viser også at selv om man utvider gruppen land til å inkludere flere, endrer det i liten grad på konklusjonene.

⁵ Merk at produktet av en matrise og en vektor er en vektor.

Basert på valgene av tidsperiode, egenskaper å matche på og donorutvalg blir den syntetiske kontrollen for EU-12 og Norge gitt av vektene i henholdsvis tabell 1 og tabell 2. Det syntetiske EU-12 og det syntetiske Norge har egenskapene som er oppgitt i hhv. tabell 3 og tabell 4 før behandling, dvs før 1. januar 1993 for EU-12 og før 1. januar 1994 for Norge.

Donorland	Vekt
Australia	0,01
Canada	0,00
Israel	0,37
Japan	0,15
New Zealand	0,00
USA	0,47

Tabell 1. Sammensetning av syntetisk kontrollgruppe for EU-12 i Lehtimäki og Sondermann (2022). Vektene summerer seg til 1.

Donorland	Vekt
Australia	0,00
Canada	0,362
Israel	0,638
Japan	0,00
New Zealand	0,00
USA	0,00

Tabell 2. Sammensetning av syntetisk kontrollgruppe for Norge. Vektene summerer seg til 1. Kilde: Jonne Lehtimäki.

Tabell 1 viser vektene som er valgt for det syntetiske EU-12. USA, Israel og Japan har høyest vekt, mens Australia står for en brøkdel. New Zealand og Canada har null vekt.

Tabell 2 viser vektene som er valgt for det syntetiske Norge. Australia, Japan, New Zealand og USA har null vekt, mens Israel står for 2/3 og Canada for 1/3.

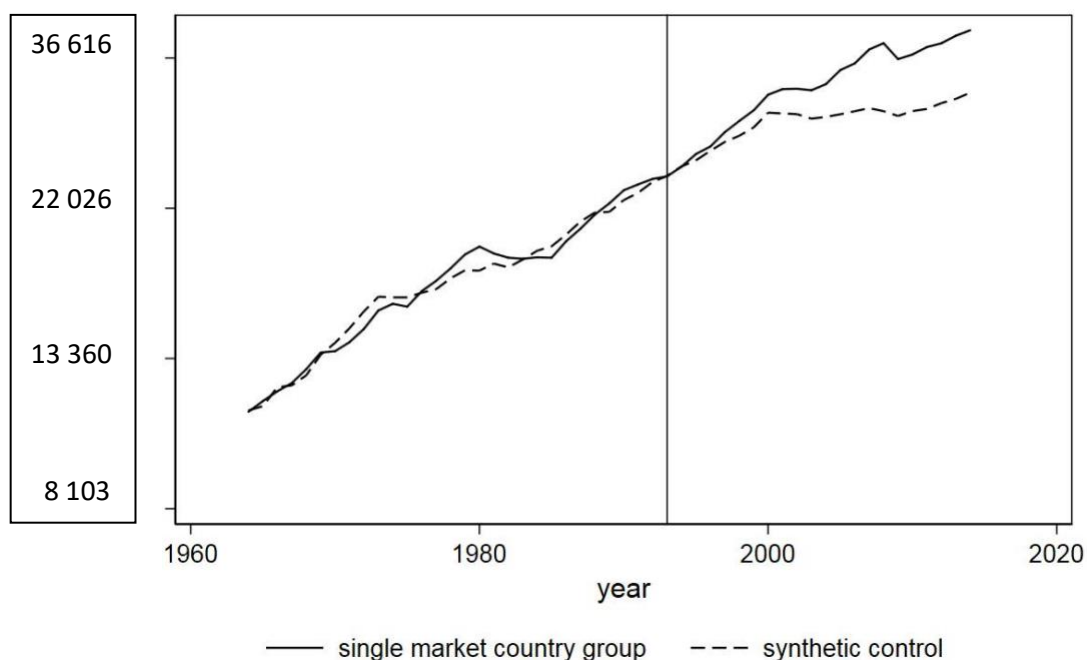
Variabel	Variabelverdier EU-12	Variabelverdier syntetisk EU-12
Logaritmen av BNP per capita i utgangsåret	9,32	9,49
Totalfaktorproduktivitet	0,82	0,84
Humankapital	2,65	3,04
Kapitalstokk	11,29	10,91
Befolkningsendring	1,48	1,21
Indeks for økonomisk kompleksitet	1,67	1,41
Indeks for økonomisk frihet	6,44	6,49
Åpenhet for handel	34,82	35,95
Investering som andel av BNP	0,28	0,28

Tabell 3. Middelerverdier (vektet gjennomsnitt) for variablene som brukes for å sammenligne EU-12 med det syntetiske EU-12. Metoden tilstreber størst mulig sammenfall mellom de to.

Variabel	Variabelverdier Norge	Variabelverdier syntetisk Norge
Logaritmen av BNP per capita i utgangsåret	9,493	9,388
Totalfaktorproduktivitet	0,756	0,895
Humankapital	2,945	2,889
Kapitalstokk	11,578	11,089
Befolkningsendring	0,021	0,164
Indeks for økonomisk kompleksitet	1,147	0,871
Indeks for økonomisk frihet	6,060	5,380
Åpenhet for handel	73,940	60,835
Investering som andel av BNP	0,354	0,270

Tabell 4. Middelerverdier (vektet gjennomsnitt) for variablene som brukes for å sammenligne Norge med det syntetiske Norge. Metoden tilstreber størst mulig sammenfall mellom de to.

Hvor godt disse syntetiske kontrollgruppene matcher EU-12 og Norge langs sentrale økonomiske variable, kan man se i tabell 3 og 4. Tabell 3 viser middelerverdien på tvers av indikator og tid for den behandlede gruppen og kontrollgruppen.



Figur 25a. Effekten av det indre marked illustrert med den syntetiske kontrollmetoden. Gjennomsnittlig effekt på reelt BNP per innbygger for EU-12-landene, målt i faste euro. I

2008 var effekten 12 prosent høyere BNP enn i et syntetisk EU-12 der det indre marked ikke ble gjennomført, i 2014 22 prosent høyere. Dette er altså en anslått kumulativ effekt, der veksten hvert år er blitt noe høyere enn man ellers ville observert, slik at BNP over tid er blitt vesentlig større. Hentet fra Lehtimäki og Sondermann (2022).

For EU-12 er nivået på BNP per innbygger i utgangsåret veldig likt. Det samme gjelder for TFP, andelen investeringer i BNP, åpenhet for handel og indeksen for økonomisk frihet. Jo likere, desto bedre. Når det gjelder middelverdien av endringen i befolkning, humankapital, en indikator for økonomisk kompleksitet og kapitalbeholdning er forskjellene mellom behandlet og syntetisk gruppe noe større. Likevel må det sies å være en betydelig match mellom den behandlede enheten og den syntetiske kontrollgruppen.

For Norge er det jevnt over ok match med den syntetiske kontrollgruppen, selv om den ikke er like god som for EU-12. BNP, produktivitetsnivå og kapitalstokk ligger nære. Det syntetiske Norge er mindre åpent for handel og har en lavere investeringsandel i BNP. Den største forskjellen gjelder befolkningsvekst i perioden. I perioden 1964–1994 var det nokså moderat befolkningsvekst i Norge (+17 prosent), mens den var høy i Canada (+50 prosent) og svært høy i Israel (+118 prosent). Langs denne variabelen er det altså dårlig samsvar før behandlingsperioden (som er det avgjørende her), men også etter 1994 var befolkningsveksten høyere i Canada og Israel. Dette avviker er verdt å være obs på, og bidrar til at de høye resultatene for Norge fra Lehtimäki må tolkes med varsomhet. Samtidig er det rimelig å konkludere med at det samlet sett er godt samsvar mellom variabelverdiene for Norge og det syntetiske Norge.

Forskjeller-i-forskjeller-metoden

En klassisk forskjeller-i-forskjeller-analyse⁶ kan sammenlignes med resultatene fra den syntetiske kontrollmetoden.

I en standard forskjeller-i-forskjeller-analyse sammenlignes forskjellen i BNP før og etter opprettelsen av det indre markedet i land som blir med i det indre markedet, med forskjellen i BNP over samme tidsperiode i land som ikke blir med i det indre markedet (kontrollgruppen).

Sammenligningen over tid, før og etter opprettelsen av det indre markedet, gjør at man tar hensyn til eventuelle nivåforskjeller i BNP mellom land som blir medlem av det indre markedet og kontrollgruppen, altså forskjeller som ikke skyldes opprettelsen av det indre markedet og som er konstante over tid. Sammenligningen mellom land som blir med i det indre markedet og kontrollgruppen gjør at man tar hensyn til endringer i BNP over tid som ikke skyldes opprettelsen av det indre markedet, som for eksempel generell teknologisk utvikling. Til forskjell fra den syntetiske kontrollmetoden består kontrollgruppen av ett eller flere land, og ikke et vektet gjennomsnitt av land.

⁶ På engelsk differences-in-differences analysis.

Da estimerer man ligningen

$$y_{t,i} = \alpha_t + \beta_1 inSM_i + \beta_2 T1992_t + \beta_3 inSM_i \cdot T1992_t + u_t,$$

Her er y logaritmen av BNP per capita. Den forklares av to dummyvariabler ($inSM_i$ og $T1992_t$). Den første er 1 for alle landene som ble med i det indre markedet fra starten av og 0 for kontrollgruppen. Den andre er 1 for alle land fra 1992 og utover og 0 før det. Det er den tredje koeffisienten som er av interesse – det er interaksjonsleddet mellom de to. Hvis fortegnet på denne er positivt og signifikant indikerer det en positiv veksteffekt på BNP per capita av å ha blitt med i det indre markedet i 1992.

	Diff-in-Diff estimate	Standard error	R^2	Observations
Baseline approach	0.138***	(0.045)	0.50	1035
With initial GDP p.c.	0.138***	(0.035)	0.69	1035
With all covariates	0.09***	(0.016)	0.93	896

Hos Lehtimäki og Sondermann (2022) gir denne modellen en veksteffekt mellom 9 og 13 prosent av det indre marked, avhengig av modellspesifikasjon. Denne kontrollsjekken konkluderer altså med en veksteffekt som kan sies å være i en sammenlignbar størrelsesorden som den syntetiske kontroll-analysen. Begge er imidlertid noe høyere enn en del resultater bygget på gravitasjons-rammeverket.