

## **Effektevaluering av Ny GIV - foreløpige resultater**

Lars Kirkebøen (SSB), Marte Rønning (SSB), Edwin  
Leuven (UiO), Oddbjørn Raaum (Frischsenteret),  
Gaute Eielsen (SSB)

Evalueringsseminar, 30. november 2012

## Program

- Innledning:
  - Om evaluering og om effekter
- Foreløpige resultater
  - Forskjeller mellom Ny GIV- og andre elever/skoler
    - ♦ Grunnskolekarakterer
    - ♦ Overgang til Vg1
    - ♦ Øvrige kjennetegn
  - Forskjeller i utvikling for Ny GIV- og andre skoler
  - Forskjeller mellom Ny GIV-deltagere og andre innen Ny GIV-skoler
- Veien videre, spørsmål

## Effektevaluering av Ny GIV

- Ny GIV-deltagere presterer dårligere enn andre elever
  - Både før og etter Ny GIV-deltagelse
- Effekt: Forskjellen mellom utfall elevene oppnår (faktisk), og hva de ville oppnådd i fravær av Ny GIV (kontrafaktisk)
  - Faktiske resultater: Lett å observere
  - Kontrafaktisk: Grunnleggende uobserverbart, må anslås

## Estimering av effekter av Ny GIV

- For å anslå hva Ny GIV-elevene ville oppnådd (kontrafaktisk) må vi ha en relevant sammenligningsgruppe
- Krever at vi har kontroll på hvordan elever og skoler fordeles til Ny GIV-deltagelse eller ikke
  - Ideelt: Ta hånd om dette ved innføring, randomisert eksperiment
  - Reelt: Nøst opp i etterkant
- Viser seg at det er betydelig variasjon i tilordningspraksis
  - Hvordan avviker Ny GIV-elever fra andre elever?
  - Hvordan avviker Ny GIV-skoler fra andre skoler?
  - Krever omfattende dataarbeid å finne relevant sammenligning

## Foreløpige resultater

- Ny GIV skal påvirke fullføring av VG
  - fase 1-elever fullfører f.o.m. 2014
- Foreløpig har vi data for
  - Grunnskolekarakterer, fase 1 og 2
  - Overgang fra GS til Vg1, fase 1

## Ny GIV-elever, -skoler og -kommuner

Skoler:	Fase 1		Fase 2		Fase 3
Elever:	Ny GIV	Alle	Ny GIV	Alle	Alle
Snitt termin 1	2.84	3.61	2.86	3.62	3.52
Andel jenter	0.41	0.48	0.36	0.49	0.48
Andel innv./barn av innv.	0.27	0.17	0.14	0.10	0.11
Mor/far høyere utdanning	0.24	0.46	0.23	0.45	0.43
Antall elever per skole	107	139	93	101	77
Antall skoler i kommunen	24.8	18.2	12.4	10.5	10.2
Antall elever	4173	37762	2675	50084	42247

6

Her - og i de følgende lysbildene – ser vi på skoleårene 2010-11 og 2011-12 under ett.

Tabellen viser noen gjennomsnittlige kjennetegn for hhv Ny GIV-elever på fase 1-skoler, alle elever på disse skolene, tilsvarende for fase 2-skoler, samt for alle elever på fase 3-skoler.

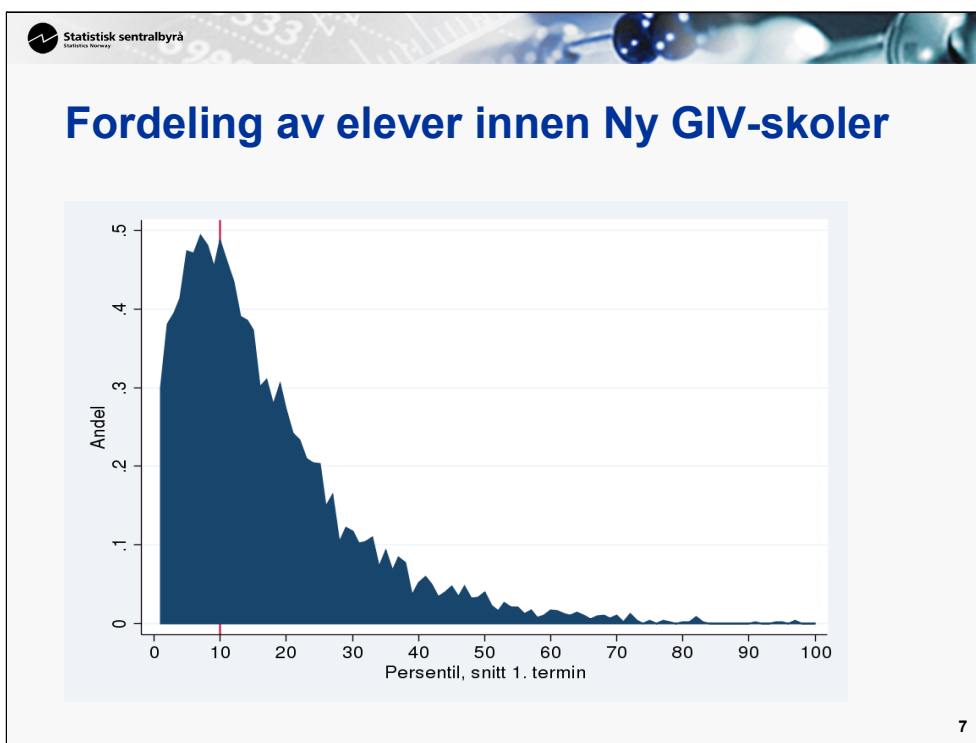
Som ventet ser vi at Ny GIV-elevene har vesentlig dårligere termin1-karakterer enn sine skolekamerater.

Videre er Ny GIV-elevene oftere gutter, innvandrere eller barn av innvandrere og de har sjeldnere høyt utdannede foreldre.

Ny GIV-skolene skiller seg først og fremst ut ved å være større skoler (med flere elever), og de ligger oftere i større kommuner med flere skoler.

Alt dette peker på at det er systematiske forskjeller mellom Ny GIV-elever og andre elever, og også mellom Ny GIV-skoler/kommuner og andre.

Dette er det viktig å ta hensyn til når vi skal vurdere effekter av Ny GIV.



Her deler vi elevene i 100 like store grupper (persentiler) innen hver kommune, sortert stigende etter snittkarakter til 1. termin (persentil 1 lavest karakter, persentil 100 høyest).

Figuren viser andel Ny GIV-deltagere per persentil i Ny GIV-skoler. Med en helt streng implementering kunne vi ventet at andelen falt fra 1 til 0 ved persentil 10.

Vi ser at det er Ny GIV-deltagere i overveiende grad er i nedre del av fordelingen, men:

- Under halvparten av elevene under 10. persentil deltar
- Andelen deltagere faller jevnt fra 10. persentil, men det finnes Ny GIV-deltagere også i øvre del

Ved 10. persentil er andelen som deltar omtrent 50/50 – hva gjør at den en elev deltar og en annen ikke?

## Sammenligning Ny GIV-elever og andre

- For å anslå hva Ny GIV-elevene ville oppnådd uten Ny GIV kan vi sammenligne deres resultater med andre elever
- Vi vet at Ny GIV-elever og -skoler i utgangspunktet er annerledes enn andre
- For at sammenligningene skal gi et bilde av effekten av Ny GIV må vi ta hensyn til alle relevante forskjeller
  - Vi kan ta hensyn til forhold som kjønn, foreldres utdanning, resultater til 1. termin...
  - Hva med andre forhold vi ikke observerer: Motivasjon, utvikling...



## Sammenligning Ny GIV-elever og andre

	Forskjell skriftlig eksamen			
Ny GIV 2011	-0.98*** (0.03)	-0.72*** (0.02)	-0.01 (0.02)	-0.27*** (0.02)
Ny GIV 2012	-1.06*** (0.02)	-0.81*** (0.02)	-0.03 (0.01)	-0.36*** (0.01)
<i>Kontroll for</i>				
Elevkjennetegn		x	x	x
Termin 1			x	
Nasj. prøver				x
Antall elever	119113	118955	118955	118955
R2	0.099	0.244	0.513	0.502

9

I første kolonne sammenligner vi Ny GIV-elevers skriftlig eksamenskarakterer med andre elevers, uten å ta hensyn til noen forskjeller. Ny GIV-elevene presterer omtrent et karakterpoeng dårligere.

Når vi tar hensyn til noen forhold som kjønn, innvandringsbakgrunn og foreldres utdanning i kolonne to reduseres forskjellen noe, men ikke veldig mye.

Når vi tar hensyn til resultater fra termin 1 blir forskjellene svært små, og ikke statistisk signifikante.

Hvis vi derimot forsøker å fange elevenes utgangspunkt ved å hensyn til resultater fra Nasjonale prøver (8. trinn) reduseres forskjellene relativt til kolonne 1 og 2, men er fortsatt ikke ubetydelige.

Vi ser altså at Ny GIV-elever presterer dårligere på eksamen, men omtrent likt når vi tar hensyn til termin 1.

Men dette kan vi ikke tolke som effektestimater:

-Fanger vi forskjeller som ikke avhenger av Ny GIV riktig ved å kontrollere for termin 1? Kan karaktersettingen til termin påvirkes av Ny GIV?

-Undervurderer vi forskjellene i utgangspunkt ved å kontrollere for NP?

-Hva med forhold som vi ikke kan observere (f.eks motivasjon, utvikling)? Vil betydningen av disse fanges av termin 1? Eller NP?

-Hva med elevene som ikke deltar i Ny GIV, kan de påvirkes av Ny GIV?

Statistisk sentralbyrå  
Statistical Central Bureau

## Elever og effekter

Skole D – deltar



Skole I – deltar ikke



 Elev deltar  
- deltakelseseffekt

 Elev deltar ikke  
- ringvirkninger

 Elev deltar ikke  
- ingen effekt

10

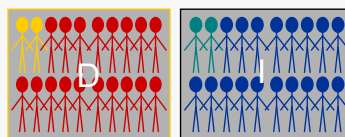
En tilnærming for å studere effekten av Ny GIV er å sammenligne skoler.

Da er det viktig å huske at også andre elever enn Ny GIV-deltagerne kan påvirkes av at skolen deltar i Ny GIV, f.eks pga endringer i organisering av undervisning.

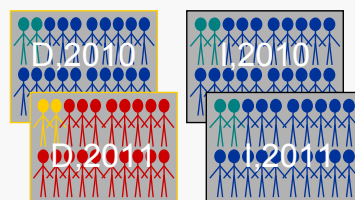
## Sammenligninger mellom skoler og år

- Kan sammenligne Ny GIV-skoler og andre skoler:

- Jevner ut uobserverbare forskjeller mellom elever?
- Effekt er sum av effekt på deltagere og andre elever



- Ny GIV-skoler og andre skoler er forskjellige - ser på forskjell i utvikling mellom skoler:



- Må anta: Uten Ny GIV utvikler skoler seg likt

- Rimelig med stabil administrasjon, lærerstab, elevgrunnlag?



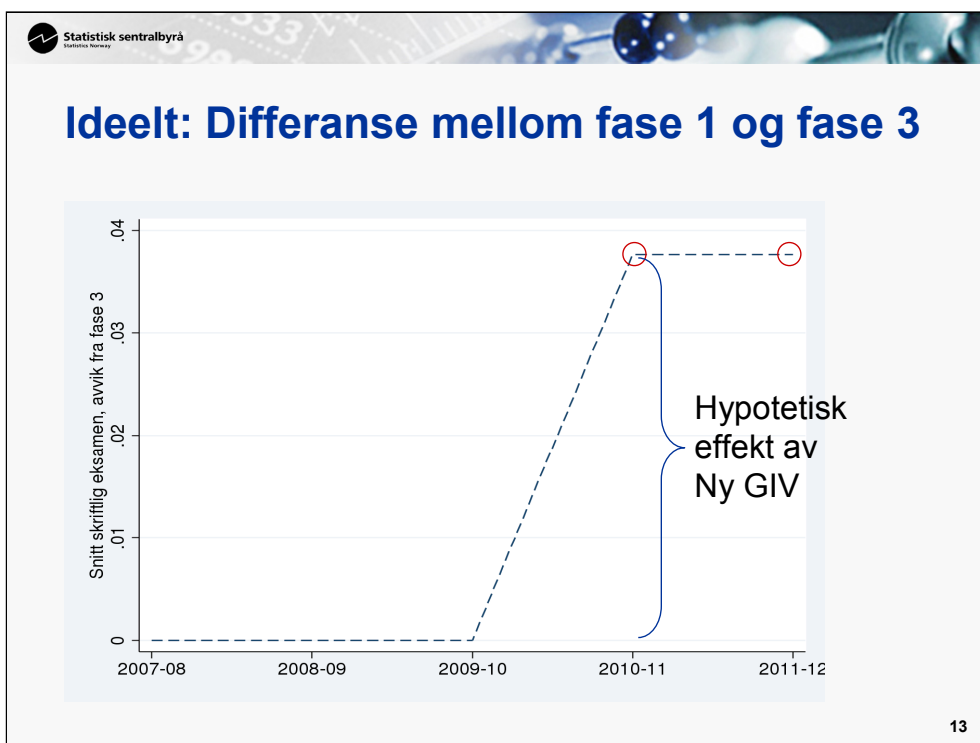
Her ser vi utviklingen i skriftlig eksamenskarakter for fase 1, 2 og 3-skoler.

Grovt sett ser vi at alle tre gruppene beveger seg på samme måte.

Grupper og år der Ny GIV er innført er ringet inn. Vi ser at eksamenskarakterene her faller fra året før.

Men det gjør de også i de samme årene for skoler uten Ny GIV.

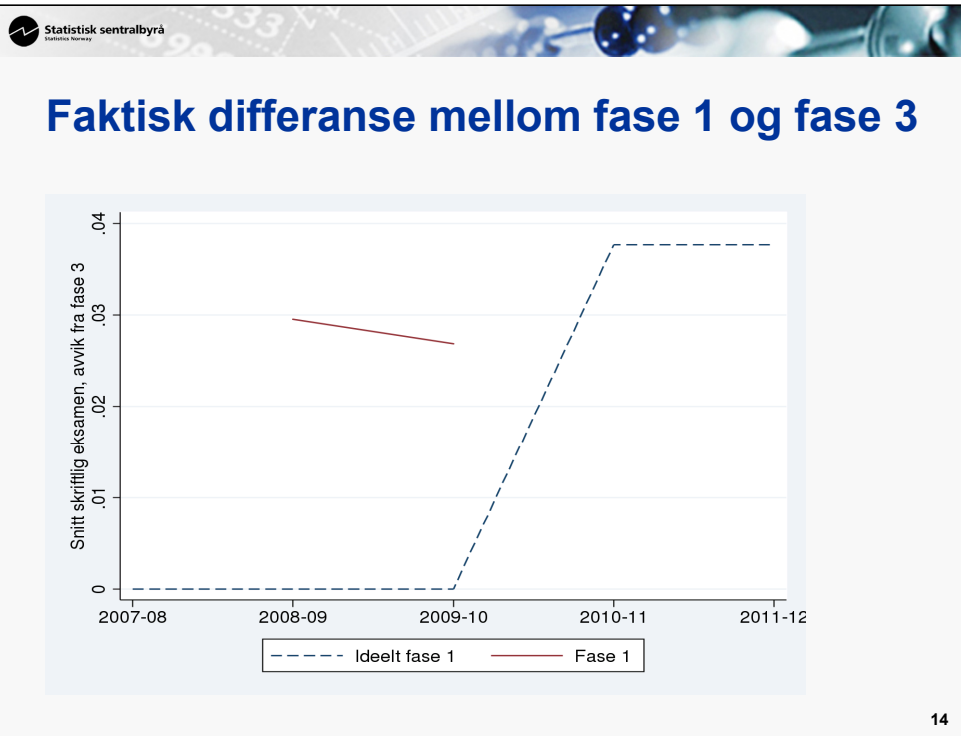
For å få et tydeligere bilde av hva som skjer bruker vi fase 3-skolene som målestokk, og går vi over til å se på differanser relativt til disse.



Før vi viser faktiske differanser: Dette er hva vi ville sett dersom Ny GIV ble innført som et ideelt eksperiment (på tilfeldige skoler), og hadde en viss stabil effekt på eksamensresultater.

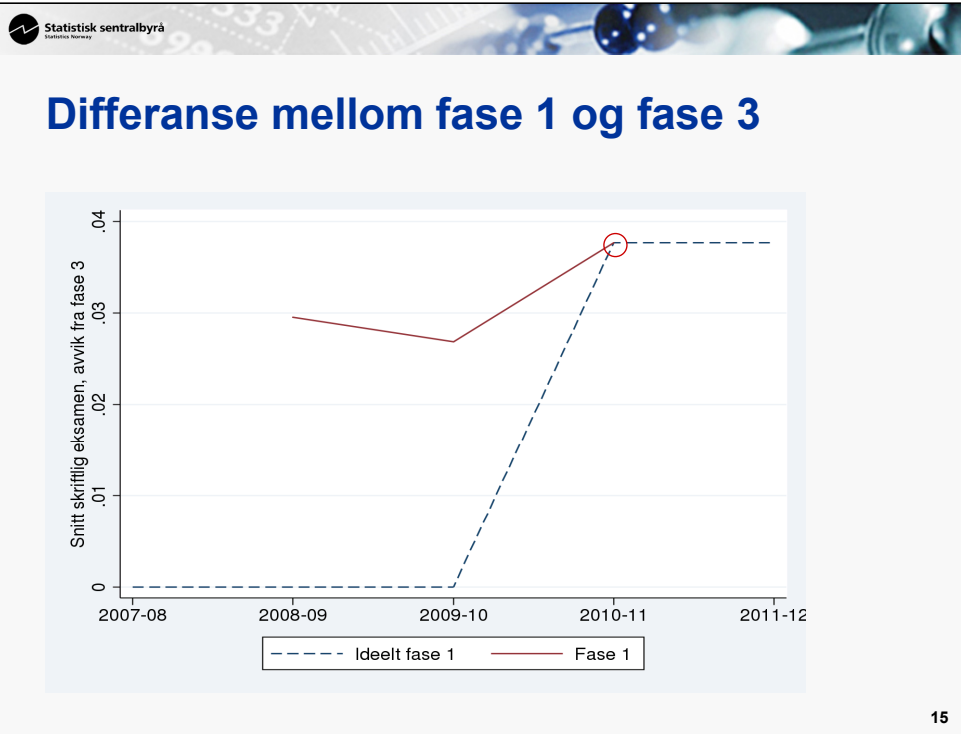
Siden skolene er valgt tilfeldig skulle resultatene vært like før Ny GIV (dvs differansen lik 0).

Ved innføring av Ny GIV ville resultatene bedret seg tilsvarende effekten (som altså ville vært et gjennomsnitt av direkte effekt og ringvirkninger).

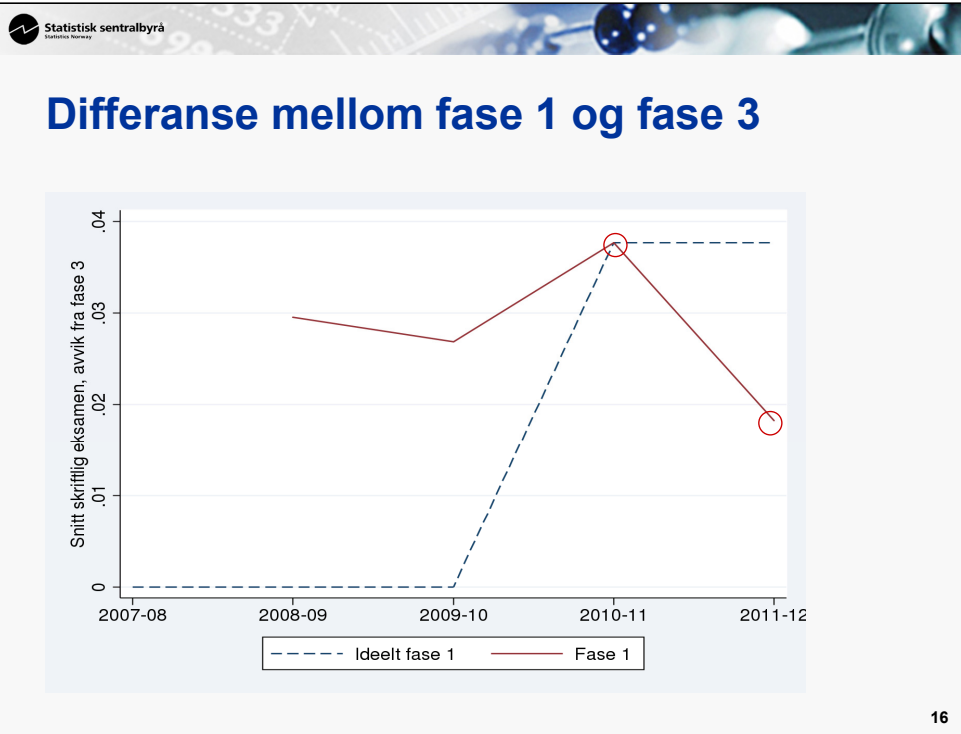


Slik skolene i fase 1 ble valgt ut avvek de systematisk fra fase 3 skolene, det ser vi også ndg eksamensresultat.

Det er imidlertid ikke noe problem, dersom avviket er stabilt over tid i fravær av Ny GIV. Dette ser ut til å være tilfellet for avgangskullene 2009 og 2010.



Så bedres eksamensresultatet såvidt, relativt til fase 3, ved innføring av Ny GIV for avgangskullet 2011.



Relativt til fase 3 faller imidlertid eksamensresultatet mer fra 2011 til 2012 enn det ble bedret fra 2010 til 2011. Dette skjer mens Ny GIV fremdeles er innført i fase 1, og ikke i fase 3.

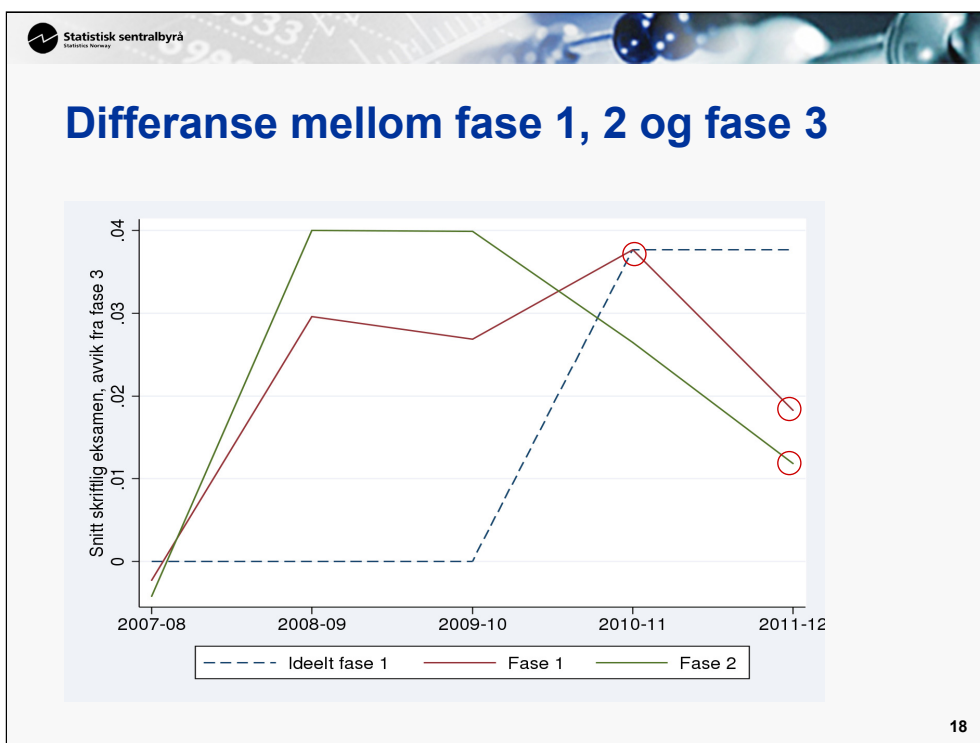


## Differanse mellom fase 1 og fase 3



17

Videre, den største endringen i forholdet mellom fase 1 og 3 skjer fra 2008 til 2009, dvs uavhengig av innføring av Ny GIV.



Dersom vi ser på fase 2-skolene, fremdeles relativt til fase 3, ser vi at eksamensresultatene faller ved innføring av Ny GIV. Men de falt tilsvarende året før, og største endring også for disse var fra 2008 til 2009.

Når vi ser at forskjellene endrer seg mer uavhengig av Ny GIV enn ifm Ny GIV har dette to mulige tolkninger:

- Alle endringene, inkl evt effekter av Ny GIV er små og ubetydelige
- Det er forskjeller ikke bare i nivå mellom gruppene, men også i utvikling, som gjør det vanskelig å si noe om effekten av Ny GIV

Det kan være vanskelig å skille mellom disse. Dette illustrerer verdien av eksperimenter, der vi ifm implementering kan sørge for at det ikke er forstyrrende forskjeller.

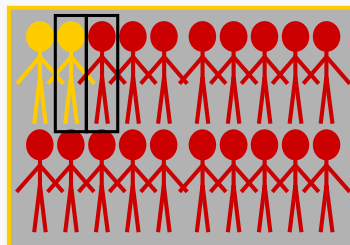


For overgang til Vg1, som vi ikke har data for 2012-kullet, ser vi at øker ved innføring av Ny GIV i fase 1, relativt til fase 3, etter å vært ganske stabil i årene før.

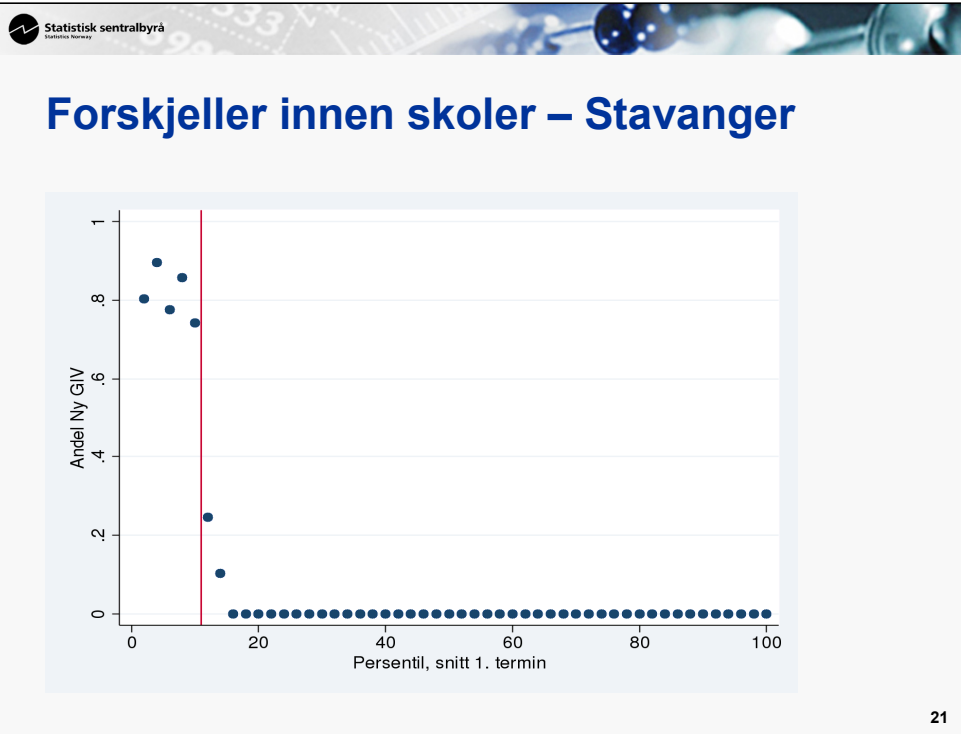
Men, samme år er økningen enda større i fase 2, som ennå ikke hadde innført Ny GIV. Dvs, vi kan ikke tolke endringen i fase 1-skolene som en effekt av Ny GIV.

## Sammenligninger innen Ny GIV-skoler

- Sammenligner elever som såvidt deltar/ikke deltar:
  - Sannsynligvis ganske like
- Krever lite antagelser – men stiller krav til implementering og data
  - I praksis vanskelig for andre enn Stavanger



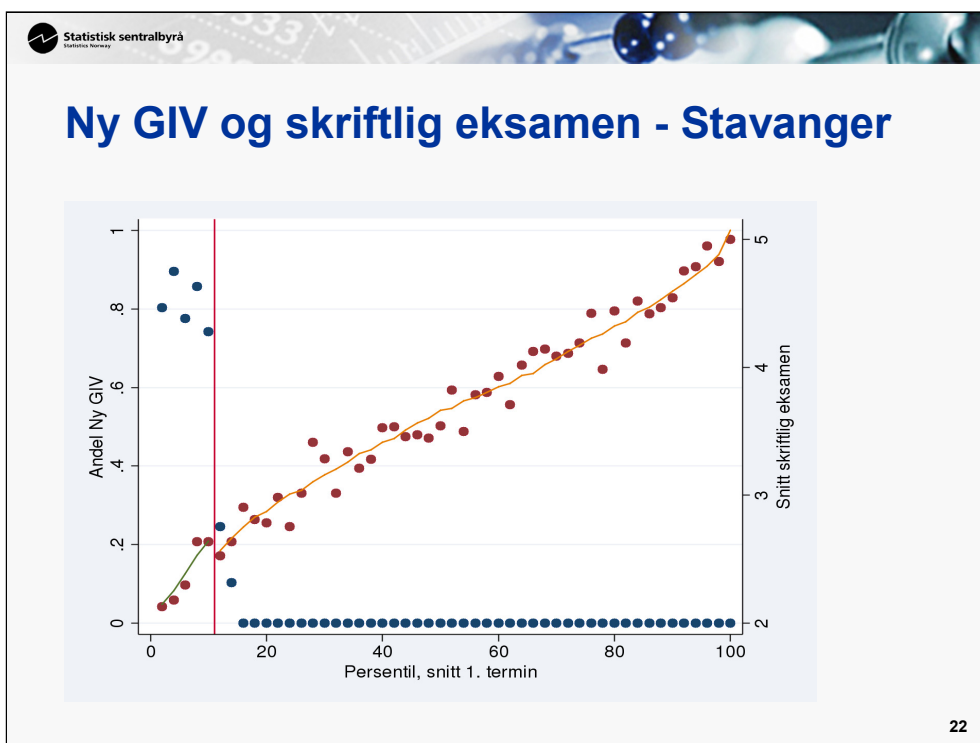
Sammenligninger innen Ny GIV-skoler er en annen tilnærming for å se på effekter av Ny GIV.



Denne figuren viser sammenhengen mellom snitt termin 1 og Ny GIV-deltagelse i Stavanger. Vi ser her bort fra elever med færre enn 11 standpunkt karakterer.

Selv om ikke alle blant de svakeste 10 prosent deltar, og det er noen over dette som deltar, der det et ganske skarpt skille i deltagelse.

Dvs, her vil elever som er ganske like, om enn ikke helt, ha svært forskjellig sannsynlighet for Ny GIV-deltagelse.



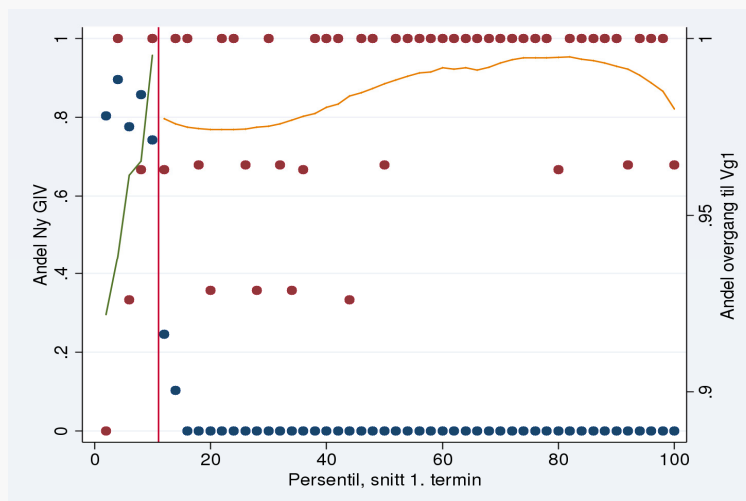
22

Vi kan vise snitt eksamenskarakter i samme diagram for Stavanger.

Dersom Ny GIV hadde en stor positiv effekt ville vi vente et fall i snittkarakteren akkurat rundt 10 prosent. Ny GIV-elever rett til venstre for den røde streken skulle da prestert bedre enn i utgangspunktet nesten like elever rett til høyre.

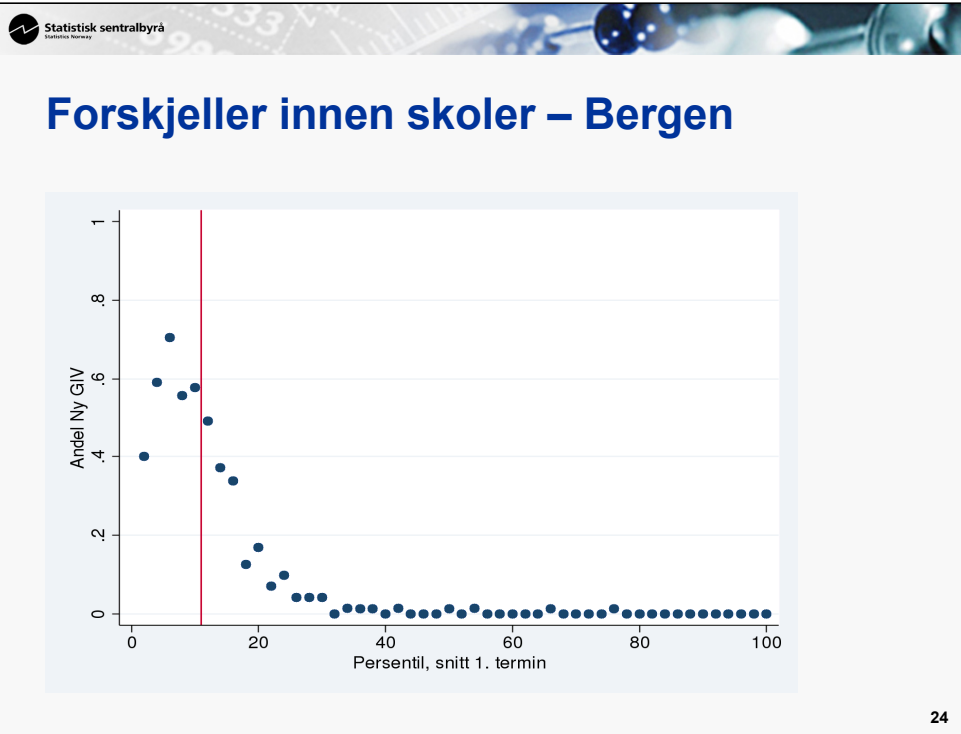
Den grønne og gule kurven viser snittkarakterer på hver side av persentil 10, til en slik sammenligning. Men, hvert punkt representerer bare omtrent 30 elever. Det betyr at usikkerheten er stor, og at vi ikke kan trekke noen klar konklusjon.

## Ny GIV og overgang til Vg1 - Stavanger



23

Tilsvarende for overgang til Vg1. Datagrunnlaget er for tynt til at vi kan trekke noen klar konklusjon.



Sammenlign fordelingen av Ny GIV-elever innen skoler f.eks med en annen stor by på Vestlandet, som kan være representativ for resten av landet.

Her er det en mer gradvis overgang, slik at det er vanskelig å sammenligne like elever som deltar/ikke deltar.



## Oppsummering

- Fremdeles tidlig – mangler data for fullføring av VG
- Ny GIV-elever og skoler er annerledes enn andre elever og skoler
  - Sannsynligvis også annerledes ndg. forhold vi ikke observerer.
- For tidlig å si noe sikkert om effekter av Ny GIV:
  - Uklart om fase 1, 2 og 3-skoler er sammenlignbare – det gir usikkerhet om hva vi finner ved å sammenligne utvikling
  - Få skoler/kommuner der det er skarpt skille mellom Ny GIV-elever og andre gir upresise resultater innen skole
  - Men: Data viser ikke tegn til store effekter på karakterer eller overgang til Vg1

## Videre

- Omfattende dataarbeid nødvendig for å konkludere:
  - Flere årganger og resultater (gjennomstrømning VG)
  - Er forskjellene i utvikling utenom Ny GIV så store at de er viktige?
  - Hvilke skoler er en relevant sammenligning med Ny GIV-skoler?
  - Er det flere kommuner som kan gi en relevant sammenligning innen skoler?
- Dette kommer vi tilbake til i 2013, og følgende år!

## Spørsmål til salen

- Lav andel Ny GIV-deltagere blant svaktpresterende?
  - Feilregistrering av deltagere?
  - Tilordning basert på annet enn snitt 1. termin?
    - ♦ Snitt av basisfag? Karaktergrenser enkeltfag?
  - IOP-elever unntatt?
  - Gjelder Ny GIV for 2011 og 2012 – strengere nå?
- Vi ser noen tegn til:
  - Færre elever med få standpunktkarakterer, og at de som har få standpunktkarakterer presterer dårligere.
  - Standpunktkarakterer svekkes, relativt til eksamen
  - Påvirker Ny GIV karaktersetting (termin 1, standpunkt)?
  - Overføres elever fra IOP til Ny GIV?

## Kontakt

- Lars Kirkebøen: [kir@ssb.no](mailto:kir@ssb.no)