



ECONnect

NTNU

Faktor

- en eksamensavis utgitt av ECONnect



Pensumsammendrag:

SØK3514 – Anvendt økonometri

Forfatter: Drago Bergholt

E-post: bergholt@stud.ntnu.no

Skrevet: Høsten 2009

Antall sider: 34



Om ECONnect:

ECONnect er en frivillig studentorganisasjon for studentene på samfunnsøkonomi- og finansøkonomistudiet ved NTNU. Vi arbeider for økt faglig kompetanse blant våre studenter samt tettere kontakt med næringslivet. Det gjør vi ved å arrangere fagdager, gjesteforelesninger, bedriftspresentasjoner m.m. I dag går det ca. 200 studenter på bachelornivå (1.-3. klasse) og ca. 70 studenter på masternivå (4.-5. klasse). Studentene på masternivå er fordelt på de to linjene samfunnsøkonomi (ca. 50 stk) og finansiell økonomi (ca. 20 stk). Mer om ECONnect og aktuelle arrangementer på www.econnect-ntnu.no.

ECONnect består av følgende personer ved utgivelsestidspunkt:

Bjørn Bergholt (Leder)	bjorn@econnect-ntnu.no
Sophie S. Strømman (Bedriftsansvarlig)	sophie@econnect-ntnu.no
Maiken Weidle (Fagdagsansvarlig)	maiken@econnect-ntnu.no
Joakim Bjørkhaug (Økonomi- og IT-ansvarlig)	joakim@econnect-ntnu.no
Elise Caspersen	elise@econnect-ntnu.no
Tiril Toftedahl	tiril@econnect-ntnu.no
Louis Dieffenthaler	louis@econnect-ntnu.no
Andreas H. Jung	andreas@econnect-ntnu.no
Mari Benedikte Ellingsen	mari@econnect-ntnu.no
Herman Westrum Thorsen	herman@econnect-ntnu.no

Post- og besøksadresse:

ECONnect, NTNU Dragvoll
 Institutt for samfunnsøkonomi
 Bygg 7, Nivå 5
 7491 Trondheim

Organisasjonsnummer:

NO 994 625 314

Hjemmeside:

www.econnect-ntnu.no

Merk: Alle pensumsammendrag og tekster som utgis av Faktor er skrevet av og for studenter. ECONnect står ikke ansvarlig for selve faginnholdet. Spørsmål om teksten kan rettes til tekstforfatteren.

SØK3514 Applied Econometrics: Summary

1. Basic Econometrics:

1.1. Ordinary Least Squares Method (OLS)

- Modell:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

Der

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + u_i \text{ og}$$

$$\bar{y} = \beta_0 + \beta_1 \bar{x}_1 + \dots + \beta_k \bar{x}_k + \bar{u}$$

- Generelle antagelser:

(1) Linearitet mellom variablene x og y i populasjonen: $y_i = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$

(2) Tilfeldig trekking fra populasjonen

(3) Variasjon i x : $Var(x) > 0$

(4) Ingen korrelasjon mellom u og x : $E(u|x) = 0$

Innebærer at $E(u) = 0$ og at $Cov(u, x) = 0$

(5) Homoskedastisitet i restleddet: $Var(u|x) = \sigma^2$

- Minste kvadraters metode - enkel regresjon:

Modell: $y_i = \beta_0 + \beta_1 x_i + u_i$ (1)

Estimat: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ (2)

Error: $\hat{u}_i \equiv y_i - \hat{y}_i$ (3)

OLS:

$$\min_{\hat{\beta}_0, \hat{\beta}_1} \sum_{i=1}^n (\hat{u}_i)^2 = \min_{\hat{\beta}_0, \hat{\beta}_1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \min_{\hat{\beta}_0, \hat{\beta}_1} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$
 (4)

(4) er problemet som løses i OLS. Førsteordensbetingelsen gir:

$$\frac{\partial \sum_{i=1}^n (\hat{u}_i)^2}{\partial \hat{\beta}_0} = 2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) (-1) = 0$$

$$\Rightarrow \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$
 (5)

$$\frac{\partial \sum_{i=1}^n (\hat{u}_i)^2}{\partial \hat{\beta}_1} = 2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) (-x_i) = 0$$

$$\Rightarrow \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0$$
 (6)

(5) og (6) kalles normalligningene. Disse brukes til å finne $\hat{\beta}_0$ og $\hat{\beta}_1$ som løser (4). Av (5):

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$
 (7)

(7) determinerer $\hat{\beta}_0$ som løser (4). Setter derfor (7) inn i (4) og minimerer med hensyn på $\hat{\beta}_1$:

$$\min_{\hat{\beta}_1} \sum_{i=1}^n (y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) - \hat{\beta}_1 x_i)^2 = \min_{\hat{\beta}_1} \sum_{i=1}^n ((y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}))^2$$
 (8)

$$\frac{\partial \sum_{i=1}^n (\hat{u}_i)^2}{\partial \hat{\beta}_1} = 2 \sum_{i=1}^n ((y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x})) (-x_i + \bar{x}) = 0$$
 (9)

$$\Rightarrow \sum_{i=1}^n \left((y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}) \right) (x_i - \bar{x}) = 0 \quad (10)$$

$$\Rightarrow \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}) - \hat{\beta}_1 \sum_{i=1}^n (x_i - \bar{x})^2 = 0$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

Oppsummert: De empiriske momentene til $\hat{\beta}_0$ og $\hat{\beta}_1$ i (2) som minimerer (4) er gitt ved (7) og (11):

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (7)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

- Forventningsrett og konsistens for $\hat{\beta}_1$:

(11) gir:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^n (\beta_1 (x_i - \bar{x}) + (u_i - \bar{u}))(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \beta_1 + \frac{\sum_{i=1}^n (u_i - \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (12)$$

Viser forventningsrett og konsistens med utgangspunkt i (12):

Forventningsrett:
$$E(\hat{\beta}_1) = \beta_1 + \frac{\sum_{i=1}^n (E(u_i|x) - \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \beta_1 \quad (13)$$

Konsistens:
$$\text{plim}_{n \rightarrow \infty}(\hat{\beta}_1) = \beta_1 + \frac{\text{Cov}(u, x)}{\text{Var}(x)} = \beta_1 \quad (14)$$

- Dekomponering av variasjonen i y :

Utgangspunkt i normallikningene (5) og (6). Disse kan omformuleres til:

$$\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = \sum_{i=1}^n \hat{u}_i = 0 \quad (15)$$

$$\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = \sum_{i=1}^n \hat{u}_i x_i = 0 \quad (16)$$

Definerer Sum of squares:

$$SST \equiv \sum_{i=1}^n (y_i - \bar{y})^2 \quad (17)$$

$$SSE \equiv \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (18)$$

$$SSR \equiv \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n \hat{u}_i^2 \quad (19)$$

$$\begin{aligned} SST &= \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n [(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})]^2 = \sum_{i=1}^n [(y_i - \hat{y}_i)^2 + 2(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) + (\hat{y}_i - \bar{y})^2] \\ &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 + 2 \sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = SSR + 2 \sum_{i=1}^n \hat{u}_i (\hat{y}_i - \bar{y}) + SSE \\ &= SSE + SSR + 2 \sum_{i=1}^n [\hat{u}_i \hat{y}_i - \hat{u}_i \bar{y}] = SSE + SSR + 2 \left(\sum_{i=1}^n \hat{u}_i \hat{y}_i - \sum_{i=1}^n \hat{u}_i \bar{y} \right) \\ &= SSE + SSR + 2 \left(\sum_{i=1}^n \hat{u}_i (\hat{\beta}_0 - \hat{\beta}_1 x_i) - \sum_{i=1}^n \hat{u}_i \bar{y} \right) \\ &= SSE + SSR + 2 \left(\sum_{i=1}^n \hat{u}_i \hat{\beta}_0 - \sum_{i=1}^n \hat{u}_i \hat{\beta}_1 x_i - \sum_{i=1}^n \hat{u}_i \bar{y} \right) = \\ &= SSE + SSR + 2 \left(\hat{\beta}_0 \sum_{i=1}^n \hat{u}_i - \hat{\beta}_1 \sum_{i=1}^n \hat{u}_i x_i - \bar{y} \sum_{i=1}^n \hat{u}_i \right) = SSE + SSR \end{aligned}$$

$$\Rightarrow SST = SSE + SSR \quad (20)$$

der kanselleringen av det siste leddet skyldes (15) og (16)

Den multiple determinasjonskoeffisienten:

$$R^2 \equiv \frac{SSE}{SST} = \frac{SST - SSR}{SST} = 1 - \frac{SSR}{SST} \quad (21)$$

- Varians og standardavvik:

Utgangspunkt i (12):

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^n (u_i - \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (12)$$

Tar variansen på begge sider:

$$\text{Var}(\hat{\beta}_1) = \text{Var}\left(\beta_1 + \frac{\sum_{i=1}^n (u_i - \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) = 0 + \frac{\sum_{i=1}^n \text{Var}(u_i | x_i)(x_i - \bar{x})^2}{[\sum_{i=1}^n (x_i - \bar{x})^2]^2} = \frac{\sum_{i=1}^n \sigma_i^2 (x_i - \bar{x})^2}{[\sum_{i=1}^n (x_i - \bar{x})^2]^2} \equiv \frac{\sum_{i=1}^n \sigma_i^2 d_i^2}{SST_x^2} \quad (22)$$

Dersom homoskedastisitet, det vil si dersom $\sigma_i^2 = \sigma_j^2 = \sigma^2$, kan (22) forenkles til:

$$\text{Var}(\hat{\beta}_1) = \frac{\sum_{i=1}^n \sigma_i^2 d_i^2}{SST_x^2} = \frac{\sigma^2 \sum_{i=1}^n d_i^2}{SST_x^2} = \frac{\sigma^2 SST_x}{SST_x^2} = \frac{\sigma^2}{SST_x} \quad (23)$$

Standardavvik:

$$Se(\hat{\beta}_1) = \sqrt{\frac{\sigma^2}{SST_x}} = \frac{\sigma}{\sqrt{SST_x}} \quad (24)$$

Empiriske momenter:

$$\text{Variansen til } u: \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n \hat{u}_i^2}{n-2} \quad (25)$$

$$\text{Variansen til } \hat{\beta}_1: \quad \widehat{\text{Var}}(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{SST_x} \quad (26)$$

$$\text{Standardavviket til } \hat{\beta}_1: \quad \widehat{Se}(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{SST_x}} \quad (27)$$

Heteroskedastisetskorrigert varians: Tar utgangspunkt i (22)

$$\widehat{\text{Var}}(\hat{\beta}_1) = \frac{\sum_{i=1}^n \hat{u}_i^2 (x_i - \bar{x})^2}{[\sum_{i=1}^n (x_i - \bar{x})^2]^2} \quad (28)$$

Heteroskedastisetskorrigert standardavvik: Tar utgangspunkt i (28)

$$\widehat{Se}(\hat{\beta}_1) = \sqrt{\frac{\sum_{i=1}^n \hat{u}_i^2 (x_i - \bar{x})^2}{[\sum_{i=1}^n (x_i - \bar{x})^2]^2}} = \frac{\sqrt{\sum_{i=1}^n \hat{u}_i^2 (x_i - \bar{x})^2}}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (29)$$

1.2. Multipl regressjon

1.3. Hypotesetesting

1.4. Heteroskedastisitet

1.5. Funksjonsform

1.6. Endogen forklaringsvariabel

1.6.1. Utelatt variabelskjevhet

- Sann modell: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$ (22)

- Estimert modell: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ (23)

- OLS-estimatoren ved enkel regresjon fant vi tidligere som (11):

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

- Setter inn for (22) der $\tilde{\beta}_1$ betegner at vi har et skjevt estimat:

$$\begin{aligned}
\tilde{\beta}_1 &= \frac{\sum_{i=1}^n (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i - \beta_0 - \beta_1 \bar{x}_1 - \beta_2 \bar{x}_2 - \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\
&= \frac{\sum_{i=1}^n (\beta_1 (x_{i1} - \bar{x}_1) + \beta_2 (x_{i2} - \bar{x}_2) + (u_i + \bar{u}))(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
&= \beta_1 \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \beta_2 \frac{\sum_{i=1}^n (x_i - \bar{x})(x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\sum_{i=1}^n (u_i + \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
&= \beta_1 + \beta_2 \frac{\sum_{i=1}^n (x_i - \bar{x})(x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\sum_{i=1}^n (u_i + \bar{u})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\
\Rightarrow \text{plim}_{n \rightarrow \infty}(\tilde{\beta}_1) &= \beta_1 + \beta_2 \frac{\text{Cov}(x_1, x_2)}{\text{Var}(x_1)} = \beta_1 + \beta_2 \delta_1
\end{aligned}$$

- Konsekvens av å utelate variabel:

- (1) $\beta_2 > 0$ og $\delta_1 > 0 \Rightarrow$ Oppadgående skjevhet
- (2) $\beta_2 > 0$ og $\delta_1 < 0 \Rightarrow$ Nedadgående skjevhet
- (3) $\beta_2 < 0$ og $\delta_1 > 0 \Rightarrow$ Nedadgående skjevhet
- (4) $\beta_2 < 0$ og $\delta_1 < 0 \Rightarrow$ Oppadgående skjevhet
- (5) $\beta_2 = 0$ og/eller $\delta_1 = 0 \Rightarrow$ Ingen skjevhet, bare større varians

1.6.2. Målefeil

- Utgangspunkt:

Sann modell: $y^* = \beta_0 + \beta_1 x^* + u$

Målefeil: $y = y^* + \varepsilon$

$x = x^* + e$

Spesifisering av forutsetninger:

- (1) $E(u) = E(\varepsilon) = E(e) = 0$
- (2) $\text{Var}(u) = \sigma_u^2, \text{Var}(\varepsilon) = \sigma_\varepsilon^2, \text{Var}(e) = \sigma_e^2$
- (3) $\text{Cov}(u, x) = \text{Cov}(u, \varepsilon) = \text{Cov}(u, e) = \text{Cov}(\varepsilon, e) = 0$

- Målefeil:

- Kun målefeil i venstresidevariabelen:

$$y = \beta_0 + \beta_1 x + u + \varepsilon = \beta_0 + \beta_1 x + v \quad \text{der } v = u + \varepsilon$$

- Kun målefeil i høyresidevariabelen:

$$y = \beta_0 + \beta_1(x - e) + u = \beta_0 + \beta_1 x + u - \beta_1 e = \beta_0 + \beta_1 x + v \quad \text{der } v = u - \beta_1 e$$

- Betingelse for skjevhet: Har tidligere vist at

$$\text{plim}_{n \rightarrow \infty}(\hat{\beta}_1) = \beta_1 + \frac{\text{Cov}(v, x)}{\text{Var}(x)} \quad (14)$$

i estimeringen av $y_i = \beta_0 + \beta_1 x_i + u_i$. Dersom det siste leddet i (14) forsvinner vil generelt ikke målefeil føre til skjevhet i estimatene. Det avgjørende er derfor hvorvidt $\text{Cov}(v, x) = 0$.

- Tilfelle 1: Kun målefeil i venstresidevariabelen:

$$y = \beta_0 + \beta_1 x + u + \varepsilon = \beta_0 + \beta_1 x + v$$

$$\text{der } \tilde{\beta}_1 = \beta_1 + \frac{\text{Cov}(v, x)}{\text{Var}(x)} = \beta_1 + \frac{\text{Cov}(u + \varepsilon, x)}{\text{Var}(x)}$$

Har allerede forutsatt at $\text{Cov}(u, x) = 0$. Dersom også $\text{Cov}(\varepsilon, x) = 0$, noe som må kunne sies å være en rimelig antagelse i de fleste tilfeller, er også $\text{Cov}(u + \varepsilon, x) = 0$. Da er ikke estimatorene skjeve, men restleddet har større varians:

$$\text{Var}(v) = \text{Var}(u + \varepsilon) = \sigma_u^2 + \sigma_\varepsilon^2 > \sigma_u^2$$

- Tilfelle 2: Kun målefeil i høyresidevariabelen; målefeil ukorrelert med observert verdi:

$$\text{Cov}(e, x) = 0$$

$$y = \beta_0 + \beta_1 x + u - \beta_1 e = \beta_0 + \beta_1 x + v$$

$$\text{der } \tilde{\beta}_1 = \beta_1 + \frac{\text{Cov}(v, x)}{\text{Var}(x)} = \beta_1 + \frac{\text{Cov}(u - \beta_1 e, x)}{\text{Var}(x)}$$

Har allerede forutsatt at $\text{Cov}(u, x) = 0$. Siden vi her forutsetter at målefeilen e er ukorrelert med observert verdi x , må $\text{Cov}(-\beta_1 e, x) = 0$. Da er ikke estimatorene skjeve, men restleddet har også her større varians:

$$\text{Var}(v) = \text{Var}(u - \beta_1 e) = \sigma_u^2 + \beta_1^2 \sigma_e^2 > \sigma_u^2$$

- Tilfelle 3: Kun målefeil i høyresidevariabelen; målefeil ukorrelert med sann verdi:

$$\text{Cov}(e, x^*) = 0$$

$$y = \beta_0 + \beta_1 x + u - \beta_1 e = \beta_0 + \beta_1 x + v$$

$$\text{der } \tilde{\beta}_1 = \beta_1 + \frac{\text{Cov}(v, x)}{\text{Var}(x)} = \beta_1 + \frac{\text{Cov}(u - \beta_1 e, x)}{\text{Var}(x)}$$

Har forutsatt at $\text{Cov}(u, x) = 0$. Men, siden $\text{Cov}(e, x^*) = 0$ er:

$$\text{Cov}(u - \beta_1 e, x) = \text{Cov}(u - \beta_1 e, x^* + e) = -\beta_1 \sigma_e^2$$

Videre har vi at:

$$\text{Var}(x) = \text{Var}(x^* + e) = \sigma_{x^*}^2 + \sigma_e^2$$

Estimatet til β_1 blir dermed:

$$\tilde{\beta}_1 = \beta_1 - \frac{\beta_1 \sigma_e^2}{\sigma_{x^*}^2 + \sigma_e^2} = \beta_1 \left(1 - \frac{\sigma_e^2}{\sigma_{x^*}^2 + \sigma_e^2} \right) = \beta_1 \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_e^2} = \beta_1 \frac{1}{1 + \frac{\sigma_e^2}{\sigma_{x^*}^2}}$$

der $\frac{\sigma_e^2}{\sigma_{x^*}^2}$ betegner noise-to-signal-ratio. Ser at desto større denne brøken er, desto skjevare er estimatet av β_1 mot null.

1.7. Tidsseriedata

1.8. Seriekorrelasjon

1.9. Difference-in-difference-estimatoren

Notasjonsforklaring:

$$D_x = \begin{cases} 1 & \text{hvis behandlingsgruppe} \\ 0 & \text{hvis kontrollgruppe} \end{cases}$$

$$D_t = \begin{cases} 1 & \text{hvis tid etter behandling} \\ 0 & \text{hvis tid før behandling} \end{cases}$$

Strukturlikning:

$$y = \beta_0 + \beta_1 D_x + \beta_2 D_t + \beta_3 D_x D_t$$

Forventet verdi $E(y|x, t)$:

$$(1) \quad \text{Kontrollgruppe før behandling:} \quad E(y|x = 0, t = 0) = \beta_0$$

$$(2) \quad \text{Behandlingsgruppe før behandling:} \quad E(y|x = 1, t = 0) = \beta_0 + \beta_1$$

$$(3) \quad \text{Kontrollgruppe etter behandling:} \quad E(y|x = 0, t = 1) = \beta_0 + \beta_2$$

$$(4) \quad \text{Behandlingsgruppe etter behandling:} \quad E(y|x = 1, t = 1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$$

Difference-in-difference-estimatoren:

$$\begin{aligned} \hat{\beta}^{DID} &= [E(y|x = 1, t = 1) - E(y|x = 0, t = 1)] - [E(y|x = 1, t = 0) - E(y|x = 0, t = 0)] \\ &= [(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2)] - [(\beta_0 + \beta_1) - (\beta_0)] = \beta_3 \end{aligned}$$

$\hat{\beta}^{DID}$ er et mål på den gjennomsnittlige behandlingseffekten ATT (ekstraeffekten for behandlede i forhold til ikke-behandlede).

1.10. Paneldata

1.11. Error Correction Model:

Utgangspunkt:

$$y_{it} = \alpha_i + \rho y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} + u_{it}$$

Trekker fra y_{it-1} på begge sider:

$$y_{it} - y_{it-1} = \alpha_i + \rho y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} - y_{it-1} + u_{it}$$

$$\Rightarrow \Delta y_{it} = \alpha_i + (\rho - 1)y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} + u_{it}$$

Legger til $\beta_0(x_{it-1} - x_{it-1})$ på høyresiden:

$$\begin{aligned} \Delta y_{it} &= \alpha_i + (\rho - 1)y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} + \beta_0(x_{it-1} - x_{it-1}) + u_{it} \\ &= \alpha_i + (\rho - 1)y_{it-1} + \beta_0(x_{it} - x_{it-1}) + (\beta_0 + \beta_1)x_{it-1} + u_{it} \\ &= \alpha_i + (\rho - 1)y_{it-1} + \beta_0 \Delta x_{it} + (\beta_0 + \beta_1)x_{it-1} + u_{it} \\ &= \gamma_i + \gamma_1 y_{it-1} + \gamma_2 \Delta x_{it} + \gamma_3 x_{it-1} + u_{it} \end{aligned}$$

der:

- (1) $\gamma_i = \alpha_i$
- (2) $\gamma_1 = \rho - 1$
- (3) $\gamma_2 = \beta_0$
- (4) $\gamma_3 = \beta_0 + \beta_1$

Kortsiktig effekt: $\gamma_2 = \beta_0$

fordi endring i x gir endring i y

Overgangseffekt: $\gamma_3 = \beta_0 + \beta_1$

fordi nivået i fjor påvirker endringen i år

Langsiktig effekt: $-\frac{\gamma_3}{\gamma_1} = \frac{\beta_0 + \beta_1}{1 - \rho}$

fordi lang sikt innebærer at $\Delta x = \Delta y = 0$

$$0 = \gamma_i + \gamma_1 y_{it-1} + 0 + \gamma_3 x_{it-1} + u_{it}$$

$$\Rightarrow \gamma_1 y_{it-1} = -\gamma_i - \gamma_3 x_{it-1} - u_{it}$$

$$\Rightarrow y_{it-1} = -\frac{\gamma_i}{\gamma_1} - \frac{\gamma_3}{\gamma_1} x_{it-1} - \frac{u_{it}}{\gamma_1}$$

Problem med lag: Kan gi skjeve estimater. Motvirker dette ved å ha tilstrekkelig stor t ($t \geq 15$).

1.12. First Difference (FD)

1.13. Fixed Effects (FE)

1.14. Instrumentvariabelmetoden og 2-SLS

2. Development Economics

2.1. Acemoglu, Johnson & Robinson: The Colonial Origins of Comparative Development: An Empirical Investigation

- Problemstilling: Virker institusjoner stimulerende på lands inntektsnivå?
 - Observasjon: Land har liknende vekstrater, men svært ulikt inntektsnivå.
 - Spørsmål: Virker institusjoner på lands inntektsnivå?
- Teori: Gode institusjoner legger forholdene til rette og stimulerer til økonomisk vekst.
- Deskriptivt: Etablert positiv sammenheng mellom institusjoner og inntekt.
- Utfordringer knyttet til å etablere en kausal sammenheng:
 - Rike land har råd til bedre institusjoner; et simultanitetsproblem.
 - Mange utelatte variabler som korrelerer både med institusjoner og inntekt; et utelatt variabelproblem.
 - I tillegg til observerbare politiske variable har vi også uobserverbare kulturelle variable som kan spille inn.
 - Humankapital
 - Naturressurser
 - Geografi
 - Institusjoner måles ex post; analytikere kan tendere til å måle institusjoner i rike land som for gode. Dette medfører i så fall positiv skjevhet.
 - Målefeil som gir skjevhet mot null.
 - Tidstrend påvirker inntektsnivået og kvaliteten på institusjoner; positiv skjevhet.
- Modell:
 - Klassisk vekstmodell:

$$Y_i = A_i F(K_i, H_i)$$

$$\Rightarrow \frac{1}{N_i} Y_i = \frac{1}{N_i} A_i F(K_i, H_i) = A_i F\left(\frac{K_i}{N_i}, \frac{H_i}{N_i}\right) = A_i F(k_i, h_i)$$

$$\Rightarrow y_i = A_i F(k_i, h_i)$$
 Strukturform: $\ln y_i = \beta_0 + \beta_1 \ln A_i + \beta_2 \ln k_i + \beta_3 \ln h_i + u_i$
 $\dot{k} = sy - (n_H + n_A)k = 0$ gir k som puttes inn i $y_i = (k_i)$
 - Markiw, Romer & Weil:

$$\ln y_i = \beta_0 + \beta_1 \ln s_{k_i} + \beta_2 \ln s_{h_i} + \beta_3 \ln(n_i + g_i) + u_i$$
 - Andre modeller:
 - Urbaniseringsmodeller
 - Geografisk vekst-modeller
 - Poverty trap models
 - Problem i alle disse modellene: Endogenitet i forhold til simultanitet og utelatte variable.
 - Acemoglu:

$$\ln y_i = \mu + \alpha R_i + u_i$$
- Variabler:

$\ln y_i$ - Vekst i GDP per capita i 1995
 R_i - Mål på beskyttelse mot eksproprieringsrisiko

- Acemoglu bruker instrumentvariabler for å etablere et kausalforhold mellom institusjoner og vekst.
 - Konseptet:
 - (1) $\ln y_i = \mu + \alpha R_i + \gamma \mathbf{X}_i + \varepsilon_i$
 - (2) $R_i = \lambda_R + \beta_R C_i + \gamma_R \mathbf{X}_i + u_{Ri}$
 - (3) $C_i = \lambda_C + \beta_C S_i + \gamma_C \mathbf{X}_i + u_{Ci}$
 - (4) $S_i = \lambda_S + \beta_S \ln M_i + \gamma_S \mathbf{X}_i + u_{Si}$
 - $S_i = \lambda_S + \beta_S \ln M_i + \gamma_S \mathbf{X}_i + u_{Si}$: Godt etablert i tidligere forskning at nybyggerdødelighet M_i påvirket bosetting S_i negativt. Medier rapporterte om dødeligheten i forskjellige områder slik at nybyggere kunne styre unna.
 - $C_i = \lambda_C + \beta_C S_i + \gamma_C \mathbf{X}_i + u_{Ci}$: Skiller mellom to typer kolonier: De som hadde høy bosetting og de som var opprettet for å ekstrahere ressurser. Kun de førstnevnte som ga utvikling av tidlige institusjoner (ca. 1900) på grunn av påtrykk fra nybyggerne. Altså en positiv sammenheng mellom bosetting og tidlige institusjoner.
 - $R_i = \lambda_R + \beta_R C_i + \gamma_R \mathbf{X}_i + u_{Ri}$: Svært liten variasjon i institusjoner over tid; de med gode institusjoner tidligere har gode institusjoner også i dag. Mye av forklaringen kan kobles til sunkne kostnader. Altså en positiv sammenheng mellom tidlige institusjoner og dagens institusjoner.
 - $\ln y_i = \mu + \alpha R_i + \gamma \mathbf{X}_i + \varepsilon_i$: Gode institusjoner påvirker inntekt positivt.
 - Hvorfor ikke bruke S_i eller C_i som instrument for R_i ? Mulig korrelasjon mellom disse og ε_i i strukturlikningen.
 - Instrumentvariabelmetoden:
 - Strukturlikningen:
 $\ln y_i = \mu + \alpha R_i + \gamma \mathbf{X}_i + \varepsilon_i$
 - Redusert form-likningen:
 $R_i = \zeta + \beta \ln M_i + \delta \mathbf{X}_i + v_i$
 - Krav til instrumentet:
 - (1) Relevans: $Cov(M_i, R_i) \neq 0$
 - (2) Eksogenitet: $Cov(M_i, \varepsilon_i) = 0$
 - Konsekvensene av (1) og (2) er at:

$$\ln y_i = \mu + \alpha R_i + \gamma \mathbf{X}_i + \varepsilon_i$$

$$\Rightarrow Cov(\ln y, \ln M) = \alpha Cov(\ln M, R) + Cov(\varepsilon_i, \ln M)$$

$$\Rightarrow \alpha Cov(R, \ln M) = Cov(\ln y, \ln M) - Cov(\varepsilon_i, \ln M)$$

$$\Rightarrow \alpha = \frac{Cov(\ln y, \ln M)}{Cov(R, \ln M)} - \frac{Cov(\varepsilon_i, \ln M)}{Cov(R, \ln M)}$$
- Gitt (1) og (2) innebærer dette at $\alpha^{IV} = \frac{Cov(\ln y, \ln M)}{Cov(R, \ln M)} \neq 0$
- $H_1: \alpha^{IV} > 0$ fordi $Cov(\ln y, \ln M) < 0$ og $Cov(R, \ln M) < 0$
- Empirisk moment:

$$\hat{\alpha}^{IV} = \frac{\sum_{i=1}^n (\ln y_i - \ln \bar{y})(\ln M_i - \ln \bar{M})}{\sum_{i=1}^n (R_i - \bar{R})(\ln M_i - \ln \bar{M})}$$

- Testing av instruments relevans: $Cov(R_i, M_i) \neq 0$
 - $\hat{R}_i = \hat{\lambda} + \underbrace{0,25}_{(0,07)} C_i + \hat{\gamma} \mathbf{X}_i + \hat{u}_i$
 - $\hat{R}_i = \hat{\lambda} + \underbrace{3,50}_{(0,70)} S_i + \hat{\gamma} \mathbf{X}_i + \hat{u}_i$
 - $\hat{R}_i = \hat{\lambda} - \underbrace{0,55}_{(0,14)} \ln M_i + \hat{\gamma} \mathbf{X}_i + \hat{u}_i$ (first stage)
 - Konklusjon: Relevant instrument!
- Resultater:
 - OLS: $\ln \hat{y}_i = \hat{\mu} + \underbrace{0,50}_{(0,05)} R_i + \hat{\gamma} \mathbf{X}_i + \hat{\varepsilon}_i$
 - 2SLS: $\ln \hat{y}_i = \hat{\mu} + \underbrace{0,94}_{(0,16)} \hat{R}_i + \hat{\gamma} \mathbf{X}_i + \hat{\varepsilon}_i$

der korresponderende first stage gir:

$$\hat{R}_i = \hat{\lambda} - \underbrace{0,61}_{(0,13)} \ln M_i + \hat{\gamma} \mathbf{X}_i + \hat{u}_i$$
- Sammenligner OLS og 2SLS og ser at OLS underestimerer effekten av institusjoner på inntekt (nedadgående skjevhet), alternativt skjevhet mot null!
 - Utelatt variabelproblem: Nedadgående skjevhet ved OLS skyldes her en utelatt variabel som:
 - (1) Korrelerer negativt med inntekt og positivt med institusjoner eller
 - (2) Korrelerer positivt med inntekt og negativt med institusjoner
 - Målefeil: Skjevhet mot null på grunn av målefeil der feilen er ukorrelert med sann verdi (en rimelig antagelse under de fleste omstendigheter).
 - Simultanitet mellom inntekt og institusjoner.
- Overidentifiseringstest: Test for eksogeniteten til forskjellige instrumenter.
 - Intuisjon:
 - Sammenligne estimater med forskjellige sett av instrumenter. Gir de samme resultat?
 - Hypotese: Instrumentet har ingen uavhengig effekt på avhengig variabel; det vil si at nybyggerdødelighet har ingen uavhengig effekt på inntektsnivå.
 - Bruker egne varianter av overidentifiseringstester; finner videre styrke for instrumentet.
- Konklusjoner:
 - Artikkelen viser at:
 - Nybyggerdødelighet forutsier moderne institusjoner.
 - Predikerte moderne institusjoner er viktig for inntekt.
 - Alternativt; land med forskjellig nybyggerdødelighet har forskjellige moderne institusjoner, og dermed forskjellig inntektsnivå.

- Forskerne har funnet et nyttig instrument for institusjoner.
- Men: Ikke klart hvorfor OLS underestimerer effekten.
- Mulig kritikk:
 - Modellformulering: Hva fanges opp av R_i ? Er R_i en dekkende indeks for eksproprieringsrisiko?
 - Forståelse: Hvilke kanaler skjer effektene gjennom? Sparing? Innbyggervekst? Investeringsmiljø? Altså et stort hopp i årsaksforklaringene!
 - Data: 64 og 110 land. Er de dekkende?

2.2. Burnside & Dollar: Aid, Policies and Growth

2.3. Harding & Rattsø: Industrial Labor Productivities and Tariffs in South Africa: Identification based on Multilateral Liberalization Reform

- Problemstilling: Fører tollreduksjon til høyere produktivitet blant arbeidstakere?
- Teori:
 - Tollreduksjon medfører økt konkurranse for innenlandsk sektor.
 - Konkurransen peker i retning av kostnadskutt og budsjettdisiplin.
 - I teorien innebærer dette høyere produktivitet per arbeider.
- Estimeringsutfordringer:
 - Endogenitet knyttet til forklaringsvariabelen tollreduksjon!
 - $\frac{Y}{N}$ er et svakt mål på arbeidskraftproduktivitet hvis man sammenligner på tvers av sektorer fordi disse varierer med hensyn på kapitalintensitet.
 - Tidligere studier har brukt faktorpriser og valutakurser som instrumenter for toll. Imidlertid seriøse endogenitetsproblemer knyttet til disse.
 - Problem:
 - Modell: $y_i = f(\text{sort boks})$
 - Hva er i den sorte boksen?
 - Utdanning
 - Helse
 - Lønn
 - Konkurranse
 - Teknologi
 - Infrastruktur
 - Osv
 - Felles for disse er at de alle har endogenitetsproblemer: For eksempel bidrar utdanning til produktivitet, men produktivitet bidrar også til utdanning.
- Teoretisk grunnlag for sammenhengen mellom tollsatser og produktivitet:
 - Effekten av tollbarrierer på innenlandsk industri er ikke opplagt:

- Infant industry argument: Beskyttelse av egen industri slik at denne kan vokse seg sterk og bli overlevelsedyktig i møte med internasjonal konkurranse.
 - Outward orientation argument: Internasjonal konkurranse skaper effektiv produksjon, naturlig seleksjon av sunne foretak og spillover-effekter fra internasjonale markedsledere.
 - De to argumentene peker i hver sin retning i forhold til arbeidskraftproduktiviteten.
- Instrumenttilnærming: Harding og Rattsø bruker utenlandske tollsatser som instrument for innenlandske tollsatser. De forutsetter og tester at:
 - Relevans: Utenlandske tollsatser er korrelert med innenlandske tollsatser.
 - Eksogenitet: Utenlandske tollsatser ikke korrelert med produktiviteten i industrien i Sør Afrika (direkte).
- Modellering:
 - Benchmark: $\Delta \ln y_{it} = \alpha_i + \alpha_t + \beta tr_{it-1}$ (1)
 - Utvidelse: $\Delta \ln y_{it} = \alpha + \beta tr_{it-1} + \gamma \ln y_{it-2}^* + \delta tr_{it-1} \ln y_{it-2}^*$ (2)
 - Notasjonsforklaring:
 - y_{it} - Produksjon-sysselsettingstraten for sektor i på tidspunkt t
 - y_{it}^* - Produksjon-sysselsettingstraten for sektor i på tidspunkt t i USA
 - tr_{it} - Tollrate for sektor i på tidspunkt t
 - Bekymring 1: Internasjonale sjokk som skaper tollendringer både i Sør-Afrika og i resten av verden. Kontrollerer for dette ved hjelp av tidsdummyer og sektorspesifikke dummyer for sjokk.
 - Bekymring 2: Tariffendringer i andre land kan påvirke arbeidskraftproduktiviteten direkte via eksport. Kontrollerer for eksport.
 - Kontrollerer også for verdens teknologifront y^*
- First stage:

$$tr_{it-1} = \beta_0 + \beta_1 tr_{it-2} LAC + \beta_2 tr_{it-2} MENA + \beta_3 tr_{it-2} SA + \beta_4 tr_{it-3} LAC + \beta_5 tr_{it-3} MENA + \beta_6 tr_{it-3} SA + v_{it-1}$$

Der den enkelte β hadde ca. $\hat{\beta} = 0,3$. Harding og Rattsø argumenterer for at $\hat{\beta}$ er signifikante, det vil si at instrumentene er relevante. Merk at ikke alle er det!
- Resultater:
 - OLS:
 - $\Delta \ln \hat{y}_{it} = \hat{\alpha}_i + \hat{\alpha}_t - \underbrace{0,073}_{(0,129)} tr_{it-1}$
 - $\Delta \ln \hat{y}_{it} = \hat{\alpha} - \underbrace{0,228}_{(0,153)} tr_{it-1} + \underbrace{0,068}_{(0,039)} \ln y_{it-2}^* - \underbrace{0,169}_{(0,092)} tr_{it-1} \ln y_{it-2}^*$
 - IV-2SLS:
 - $\Delta \ln \hat{y}_{it} = \hat{\alpha}_i + \hat{\alpha}_t - \underbrace{0,840}_{(0,419)} tr_{it-1}$
 - $\Delta \ln \hat{y}_{it} = \hat{\alpha} - \underbrace{0,830}_{(0,412)} tr_{it-1} + \underbrace{0,082}_{(0,077)} \ln y_{it-2}^* - \underbrace{0,194}_{(0,214)} tr_{it-1} \ln y_{it-2}^*$

- Forståelse: En tiprosentpoengs økning i tollsatsene reduserer produktivitetsveksten med nesten ett prosentpoeng.
- Viktig: OLS gir oppadgående skjevhet. Intuisjonen er at dette skyldes beskyttelse av sektorer som i utgangspunktet er mindre produktive, slik at man undervurderer effekten av tollsatter ved OLS.
- Error correction modell:
 - Teoretisk utgangspunkt:

$$\Delta y_{it} = \gamma_i + \gamma_1 y_{it-1} + \gamma_2 \Delta x_{it} + \gamma_3 x_{it-1} + u_{it}$$
 der:
 - (1) Kortsiktig effekt: γ_2 fordi endring i x gir endring i y
 - (2) Overgangseffekt: γ_3 fordi nivået i fjor påvirker endringen i år
 - (3) Langsiktig effekt: $-\frac{\gamma_3}{\gamma_1}$ fordi lang sikt innebærer at $\Delta x = \Delta y = 0$
 - Resultater:

$$\Delta \ln \hat{y}_{it} = \hat{\gamma}_i - \underbrace{0,162}_{(0,033)} \ln y_{it-1} - \underbrace{0,33}_{(0,387)} \Delta tr_{it} - \underbrace{0,746}_{(0,264)} tr_{it-1} + u_{it}$$
 - (1) Kortsiktig effekt: $\gamma_2 = -0,33$ og ikke signifikant; ingen kortsiktig effekt
 - (2) Overgangseffekt: $\gamma_3 = -0,746$
 - (3) Langsiktig effekt: $-\frac{\gamma_3}{\gamma_1} = -\frac{-0,746}{-0,162} = -4,605$
 Forståelse: En økning på 10% fra 14% til 15,4% gir: $-4,605 \times 0,14 = 0,645$
 Altså gir en 10% økning i tollsatsene fra 14% til 15,4% en reduksjon i produktivitetsveksten på 0,645%.
- Konklusjoner:
 - Omgår endogenitetsproblemet knyttet til tollsatsenes effekt på produktivitetsveksten ved hjelp av instrumentet utenlandske tollsatter.
 - Finner at nedbygging av tollbarrierer har bidratt til økt arbeidskraftproduktivtetsvekst i Sør-Afrika.
 - Resultatene gir støtte til "outward orientation argument" på bekostning av "infant industry argument".
 - Vanskelig å si hvorvidt fordelene ved handelsliberalisering er knyttet til spillover av teknologi eller til økt konkurransepress. Antageligvis har begge forhold betydning.

2.4. Björkman & Svensson: Power to the People: Evidence from a randomized field experiment on community-based monitoring in Uganda

- Problemstilling: Fører økt sivilt engasjement og overvåking til bedre offentlig administrering?
- Aktualiserende:
 - Rundt 11 millioner barn under 5 år dør hvert år, og nesten halvparten av disse lever i Afrika sør for Sahara.
 - Mulig medvirkende årsak: Dårlig utviklede overvåkningssystemer og svake ansvarsforhold i forhold til helsetjenester.
- Eksperiment:

- 50 legesentre deltok i eksperimentet.
- 25 i behandlingsgruppen og 25 i kontrollgruppen.
- Stratifisering etter geografisk lokalisering og populasjonsstørrelse.
- Diff-in-diff-estimatorene:
 - Tiltak for å sikre at den estimerte effekten ikke fanger opp noe uforklart.
 - Poeng: Hvis det er forskjeller mellom behandlingsgruppe og kontrollgruppe før behandlingen vil vanlig OLS gi skjeve estimater! Diff-in-diff-estimering kontrollerer for initiale forskjeller.
 - Diff-in-diff-estimering:

$$y = \beta_0 + \beta_1 D_x + \beta_2 D_t + \beta_3 D_x D_t$$
 Forventet verdi $E(y|x, t)$ av eksperimentet:
 - (1) Kontrollgruppe før behandling: $E(y|x = 0, t = 0) = \beta_0$
 - (2) Behandlingsgruppe før behandling: $E(y|x = 1, t = 0) = \beta_0 + \beta_1$
 - (3) Kontrollgruppe etter behandling: $E(y|x = 0, t = 1) = \beta_0 + \beta_2$
 - (4) Behandlingsgruppe etter behandling: $E(y|x = 1, t = 1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$
 - Difference-in-difference-estimatorene $\hat{\beta}^{DID} = \beta_3$ er et mål på den gjennomsnittlige behandlingseffekten ATT av å ha "community-based monitoring" av offentlig primærhelsetjeneste.
- Resultater: $\hat{\beta}^{DID}$ positiv og signifikant, der tallverdien var avhengig av definisjonen på y (ventetid, service, medisinsk tilbud, kompetanse, oppfølging, etc.).
- Kritik:
 - Fra Ratsø: "Noe vil alltid skje når man innfører et program!"
 - Dessuten: Mulig ekstraeffekt av at vestlige observatører var til stede og kontrollerte prosessen på grunn av forventningspress på lokalt helsepersonell.

3. Public Economics

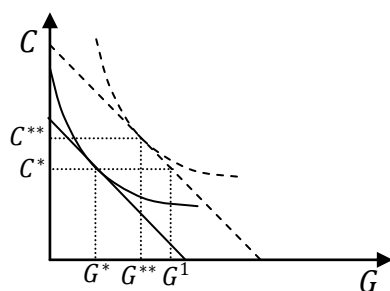
3.1. Gordon: Do federal grants boost school spending? Evidence from Title I

- Grunnskolen finansieres ved hjelp av eiendomsskatt og overføringer. Title I er et forsøk på å øke ressursene i skoledistrikter med svak sosioøkonomisk demografi.
- To konkurrerende modeller for den finansielle effekten av Title I:

- Medianvelgermodellen:

- Medianpreferanser: $U^m = (C^m, G)$
der $U_C^m > 0$ og $U_G^m > 0$

- Grafisk:



- Hvis medianvelgermodellen samsvarer med realiteten vil en økning i skoleinntektene gjennom Title I medføre økning i utgifter til både konsum og skole. Dette innebærer at økningen i skoleutgifter må være mindre enn tilskuddet. Det er mulig for myndighetene å tilsløre dette ved å la tilskuddet i sin helhet gå til skole, men heller kutte i generelle skoleutgifter.

- Inntektselastisiteten til skoletjenester: $\varepsilon \equiv \frac{\frac{dG}{G}}{\frac{dY}{Y}} = \frac{dG}{dY} \frac{Y}{G}$

- The Flypaper Effect:

- Måltrettede tilskudd har større inntektselastisitet (i forhold til målet) enn generelle tilskudd; "money sticks where it hits".
- Dersom det er sterk flypaper-effekt vil β_1 være nærme 1 i modellen
 $UNDERSVINGNSUTGIFTER = \beta_0 + \beta_1 TITLE I$

- Tidligere studier har vist at det ikke er en signifikant effekt av Title I. Men; dette skyldes minst én av de følgende to. Tjenestene finansiert gjennom Title I er:

- Ineffektive
- Ikke representative for nettoøkningen i finansiering på grunn av Title I

- Konvensjonell OLS-modell: $\Delta UNDERSVINGNSUTGIFTER = \beta_0 + \beta_1 \Delta TITLE I$

- Problemet med denne tilnærmingen er at endringen i Title I-tilskudd baseres på karakteristika ved skoledistrikter. Dermed har vi mulige utelatt variabel-problemer og simultanitet.
- Eksempel: Tilstrømming av innvandrere kan både øke undervisningsutgifter generelt og Title I-tilskudd. Dette vil i så fall gi oppadgående skjevhet.
- Løsning: Et instrument som identifiserer den eksogene variasjonen i Title I-tilskudd.

- Identifikasjonsstrategi: Title I-tilskudd ble beregnet på grunnlag av tiårige folketellinger. Gordon benytter overgangen fra 1980-census til 1990-census for å identifisere den faktiske endringen i Title I. Endring i Title I-tilskudd i 1992.
- Modellering:
 - Second stage:

$$\Delta INSTEXP_{PPd} = \beta_0 + \beta_1 \Delta \widehat{TIPP}_d + \beta_2 \Delta STATEREVPP_d + \beta_3 \Delta LOCREVPP_d + \beta_4 \Delta ENROLLMENT_d + u_d$$
 - First stage:

$$\Delta TIPP_d = \alpha_0 + \alpha_1 SIMTIPP_d + \alpha_2 \Delta STATEREVPP_d + \alpha_3 \Delta LOCREVPP_d + \alpha_4 \Delta ENROLLMENT_d + v_d$$
- Resultater:
 - OLS:
 - 1. år: $\hat{\beta}_1 = \frac{0,994}{(0,473)}$
 - 2. år: $\hat{\beta}_1 = \frac{0,955}{(0,401)}$
 - 3. år: $\hat{\beta}_1 = \frac{0,330}{(0,348)}$
 - IV-SLS:
 - 1. år: $\hat{\beta}_1 = \frac{1,401}{(0,551)}$
 - 2. år: $\hat{\beta}_1 = \frac{0,960}{(0,509)}$
 - 3. år: $\hat{\beta}_1 = \frac{0,119}{(0,501)}$
- Konklusjoner:
 - Ser ut til at vi har flypaper effect 1. og 2. året, og median voter 3. året.
 - Problem ifølge Rattsø: Endringen i Title I på grunn av folketellingen kan være ventet, og skoledistrikter kan derfor ha tilpasset seg disse endringene på forhånd.

3.2. Petterson-Lidblom: Do parties matter for economic outcomes: A regression-discontinuity approach

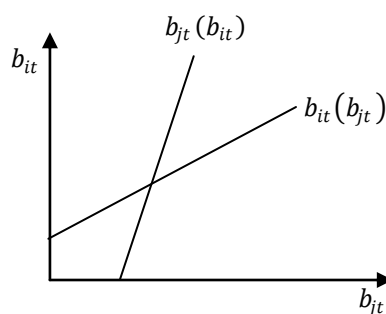
- Problemstilling: Påvirker partikontrollen i lokale styringsenheters skatte- og økonomiske politikk?
- Grunnleggende estimeringsproblem: Partier trekkes ikke tilfeldig til å lede styringsenhetene.
- Tidligere studier: $EXP_i = \beta_0 + \beta_1 SOC_i + kontroll_i + u_i$
 - Problem: Endogenitet i SOC .
 - Eksempel: Senterpartiet er store i Hedmark og Oppland, men hvordan er forbruksmønsteret i typiske landbruksregioner?

- Eksempel: Politiske sykluser der budsjettene har en tendens til å ekspandere før valg.
- Poeng: Vanskelig å ta høyde for alle forhold som korrelerer både med forbruksmønster og med typiske sosialdemokratiske områder.
- Identifikasjonsstrategi: Diskontinuitet!
 - Lokalområder med styresmakter som har stemmer/representanter rett over og under 50 % burde ha liknende karakteristika, men forskjellig politisk kontroll.
- Modellering: $\ln y_{it} = \alpha_i + \alpha_1 D_t + \alpha_2 SOC_{it} + \alpha_3 f(\text{andel venstrestemmer}) + u_{it}$
 - y_{it} - Økonomisk utfall i kommune i på tidspunkt t
 - α_i - Kommunespesifikk og tidsuavhengig variabel
 - D_t - Sett med tidsdummyer
 - SOC_{it} - Dummy lik 1 hvis styring av venstrepartier og lik null hvis styring av høyrepartier
 - $f(\cdot)$ - Polynomfunksjon av andelen venstrestemmer
 - Merk: Den avhengige variabelen er forskjellige varianter av økonomisk utfall, for eksempel utgifter per capita, skattesatser og inntekter, arbeidsledighet, antall offentlige ansatte per capita, etc.
 - Merk: Den interessante variabelen er α_2 , som måler gjennomsnittlig forskjell i økonomisk utfall mellom venstre regjeringer og høyre regjeringer.
- Resultater: Ser på forskjellen i totale utgifter per capita
 - Deskriptiv statistikk:
 - Alle:
 - Venstre: 29562
 - Høyre: 26787
 - Forskjell: 2775 (10 %)
 - ± 4 % rundt 50 %:
 - Venstre: 29237
 - Høyre: 29059
 - Forskjell: 178 (0,7 %)
 - Diskontinuitetsregresjonen:

$$\ln y_{it} = \hat{\alpha}_i + \hat{\alpha}_1 D_t + \underbrace{0,024}_{(0,009)} SOC_{it} + \hat{\alpha}_3 f(\text{andel venstrestemmer})$$
 - Altså 2,4 % større totalutgifter når venstrestyrt kommune.
 - Ser at forskjellen er signifikant.
- Mulig simultanitetsproblem: Kan partikontroll predikeres av økonomiske variable?
 - Petterson-Lidblom tester for dette og finner støtte for tilfeldig fordeling av partikontroll.
 - Finner heller ikke at partikontroll gir effekt på kontrollvariablene.
- Konklusjon:
 - Sosialdemokratisk partikontroll gir høyere offentlig utgiftsnivå.
 - Men: Signifikansen forsvinner i stor grad når diskontinuiteten kuttes til ± 2 %.

3.3. Fiva & Rattsø: Welfare competition in Norway: Norms and expenditures

- Er det velferdskonkurransen mellom kommuner? Altså; medfører økt velferdstilbud i én kommune økt velferdstilbud også i nabokommuner?
- Teori:
 - Naboeffekt: Mobilitet innebærer at det får konsekvenser hvis enkeltkommuner skiller seg vesentlig fra andre kommuner.
 - Selv uten mobilitet vil velferdstilbudet i nabokommuner gi informasjon og sammenligningsgrunnlag for velgerne i forhold til eget tilbud.
 - Eksempel:
 - Økt velferdstilbud i kommune j kan medføre fraflytting i kommune i hvis ikke denne svarer med økt velferdstilbud.
 - Dette vet kommunale styresmakter (politikere og administrasjon) i kommune i , man vil forsøke å utjevne forskjellen i velferdstilbudet.
 - Men: Kan også tenkes et "race to the bottom" fordi kommuner ønsker å tiltrekke seg skattebetalere og frastøte seg trygdebrukere. I så fall vil et "race to the bottom" innebære nedbygging av både skatteregimer og velferdstilbud. Wheaton beskriver trygdebrukere som relativt mobile sammenlignet med skattebetalere.
 - Implikasjonen av sistnevnte fremstilling er at kommunalt bestemte velferdstilbud generelt vil være mindre enn nasjonalt bestemte velferdstilbud.
 - Uansett; i begge tilfeller vil det være en positiv sammenheng mellom velferdstilbudet i nærliggende kommuner. Grafisk kan sammenhengen illustreres med stigende reaksjonskurver:



- Målet i denne artikkelen: Å estimere hvor følsomme velferdstilbudene i nabokommuner er for hverandre.
- Modell: $b_{it} = \beta_0 + \sum_{j=1}^n (w_{ij} b_{jt}) + \sum_{k=1}^K (\beta_k x_{ikt}) + x_{it}$
- Estimeringsproblem: Det følger av fremstillingen over at dette er en simultanmodell; b_{it} og b_{jt} bestemmes simultant.
- Fiva & Rattsø sammenligner tre estimeringsmetoder: IV, OLS og Maximum Likelihood.
- Resultater:
 - IV: $\underbrace{0,80}_{(0,19)}$

- OLS: $\underbrace{0,61}_{(0,08)}$
- ML: $\underbrace{0,39}_{(0,06)}$
- Alle metodene viser signifikante naboeffekter.
- Tolkning: En økning på 1000 NOK i velferdstilbudet i kommune j gir en økning på 800 NOK i velferdstilbudet i kommune i (jf. OLS). Altså en sterk effekt!
- Konklusjon: Man kan ikke forstå velferdsnivået i én kommune uten kunnskap om velferdsnivået i andre kommuner.
- Kritikk: Med nabokommuner menes de som deler grense. Imidlertid kan det hende at innbyggere sammenligner velferdstilbudet i egen kommune med velferdstilbudet i kommuner som har andre fellestrekk enn grense. For eksempel virker det rimelig at Trondheim kommune sammenlignes med Bergen heller enn med Melhus innenfor mange velferdsområder.

3.4. Borge & Rattsø: Property taxation as incentive for cost control: Empirical evidence for utility services in Norway

- Problemstilling: Noen kommuner har kommuneskatt, andre har det ikke. Har kommuneskatten noe å si for kostnadsdisiplinen?
- Teori:
 - Folk som pålegges eiendomsskatt er mer opptatt av hvordan offentlige inntekter brukes.
 - Her måles bruken av skatteinntektene her som kommunale kloakkutgifter.
- Deskriptivt: Kommuner med eiendomsskatt har ca. 20 % lavere kostnadsnivå enn de uten eiendomsskatt.
- Eiendomsskatten:
 - I motsetning til så og si all annen kommunal inntekt er tilstedeværelsen og størrelsen på eiendomsskatten kommunalt bestemt.
 - Skiller mellom eiendomsskatt rettet mot privat og kommersiell eiendom på den ene siden, og eiendomsskatt rettet mot kraftstasjoner på den andre siden.
 - Siden sistnevnte kan tolkes som eksport av skattebyrden skal vi her se på førstnevnte.
- Estimeringsproblem: Ingen tilfeldig fordeling av eiendomsskatt. Typisk regioner under sosialdemokratisk kontroll som har eiendomsskatt. Hvis sosialdemokratisk kontroll innebærer større offentlige utgifter vil dette medføre at man undervurderer effekten på kostnadskontroll av eiendomsskatt.
- Bruker to tilnærminger:
 - OLS
 - Matching OLS
- Identifikasjonsstrategi: Matching analysis
 - Steg 1:

Estimerer relasjonen

$$\widehat{PRTAX}_i = \beta_0 + \beta_1 \text{bosettingsmønster}_i + \beta_2 \text{kommunal inntekt}_i + \beta_3 \text{andel sosialister}_i + \text{etc} + v_i$$

der \widehat{PRTAX}_i er estimert sannsynlighet for at kommune i har eiendomsskatt.

○ Steg 2:

Likningen over er en lineær sannsynlighetsmodell (binær venstresidevariabel), og kan brukes til å identifisere kommuner med liknende sannsynlighet for å ha eiendomsskatt. Stratifiseringen må være slik at kommunene innenfor hver klasse (block) ikke har signifikant forskjellig sannsynlighet for eiendomsskatt. I denne undersøkelsen krevde dette 5 klasser.

○ Steg 3:

Estimerer relasjonen

$$\ln c_{it} = \beta_t + \beta_1 \widehat{PRTAX}_i + u_{it}$$

innenfor hver klasse, og tester hvorvidt $\beta_1 < 0$.

• Resultat:

- OLS: Finner at $\beta_1 = -0,10$ og signifikant.
- Matching OLS: Finner at $\beta_1 = -0,17$ og signifikant.
- Forståelse: Kommuner med eiendomsskatt har mellom 10 % og 17 % lavere kostnadsnivå enn kommuner uten eiendomsskatt.

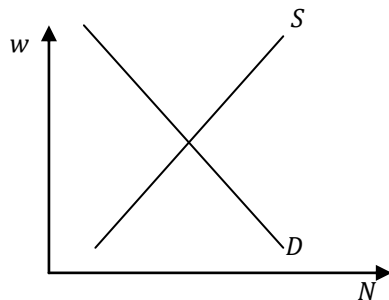
• Robusthetstest: Har kommuner med betydelige fiskale problemer oftere eiendomsskatt?
Nei.

• Konklusjon: Ser ut som om eiendomsskatt kan være en insentivmekanisme for lavere offentlige utgifter.

4. Labour Economics

4.1. Benmarker, Mellander & Öckert: Do regional payroll taxes boost employment?

- Problemstilling: Har reduksjoner i arbeidsgiveravgiften positiv effekt på sysselsetting og lønn?
- Klassisk modell:
 - Antar stigende tilbudskurve og fallende etterspørselskurve.
 - Grafisk:



- Redusert inntektsskatt vil skifte tilbudskurven ut, redusert arbeidsgiveravgift vil skifte etterspørselskurven ut.
- I begge tilfeller vil noe av skattereduksjonen falle over på parten i arbeidsmarkedet som ikke møter skattereduksjonen direkte.
- I begge tilfeller vil sysselsettingen øke.
- Ved reduksjon i inntektsskatt vil lønnen gå ned, ved reduksjon i arbeidsgiveravgift vil lønnen gå opp.
- Mer realistisk i Skandinavia: Right to manage modell
 - Modellen oppsummert:
 - Prosess: Fagforeningen og bedriften fremforhandler lønn. Deretter bestemmer bedriften sysselsetting. Utfallet avhenger av partenes relative forhandlingsmakt.
 - Lavere arbeidsgiveravgift innebærer at marginalkostnaden på arbeidskraft blir lavere enn marginalproduktiviteten. Dette gir et positivt skift i etterspørselen. Samtidig vil fagforeninger benytte det faktum at gitt arbeidskraft har blitt billigere til å kreve høyere lønn.
 - I den grad fagforeninger:
 - Observerer redusert arbeidsgiveravgift på forhånd
 - Er opptatt av økt lønn fremfor økt sysselsetting
 - Har relativ forhandlingsmakt sammenlignet med bedriftene
 - Har hyppige forhandlinger med motparten
 vil lavere arbeidsgiveravgift resultere i økt lønn.
 - På den andre siden vil uventede avgiftsreduksjoner gi økt sysselsetting, i det minste frem til neste forhandlingsrunde.

- Utgangspunkt: Kommuner i nordre Sverige uten kystlinje fikk 10 prosentpoengs reduksjon i arbeidsgiveravgiften fra 38% til 28% i 2002.
- Modellering:
 - Strukturell form:

$$y_{jkrt} = \beta \tau_{jrt} + \gamma_{kt} + \lambda_j + \varepsilon_{jkrt}$$
 der y_{jkrt} er utfall (sysselsetting, lønn, etc) for bedrift j i industrisektor k i region r på tidspunkt t . τ er arbeidsgiveravgift, γ_{kt} er sektorspesifikk og tidsavhengig effekt, λ_j er bedriftsspesifikk effekt.
 Problem: Endogenitet i τ .
 Løsning: Instrumentering av τ .
 - Redusert form:

$$\hat{t}_{jkrt} = \phi D_R D_T + \delta_{kt} + \varpi_j + \eta_{jkrt}$$
 der \hat{t}_{jkrt} er estimert arbeidsgiveravgift, D_R er en dummy for regionene med avgiftskutt, D_T er en dummy for tiden etter avgiftskuttet, δ_{kt} er sektorspesifikk og tidsavhengig effekt, ϖ_j er bedriftsspesifikk effekt.
 - Forståelse av β i strukturlikningen: En diff-in-diff-estimator der man sammenligner forskjell i y mellom perioden før og etter avgiftskutt mellom bedrifter som ble truffet av kuttet og bedrifter som ble holdt utenfor.
 - Matching: Sammenligner bedrifter som fikk avgiftskutt med nærliggende bedrifter som ikke fikk avgiftskutt. Antar dermed at det ikke eksisterer noen forskjeller i underliggende trender mellom bedrifter i forskjellige regioner.
 - Dessuten: Estimerer FE-estimatoren for å ta høyde for permanente forskjeller mellom bedrifter, permanente forskjeller mellom sektorer, samt permanente forskjeller mellom regioner.
- Resultater:
 - Etablerte bedrifter (ekskudert entry/exit): Ingen signifikant sysselsettingseffekt, men signifikant lønnseffekt.
 - Alle bedrifter (inkludert entry/exit): Signifikant effekt på antall bedrifter. Dessuten nøytraliseres lønnsøkningen i eksisterende bedrifter av lav lønn i nyetablerte bedrifter.
- Kritikkk: Hvor plutselig skjedde reduksjonen i arbeidsgiveravgiften? En forventet reduksjon kan gi tilpasninger i forkant.

4.2. Falch: The elasticity of labor supply at the establishment level

- Problemstilling:
 - Bedriftens tilpasning:

$$\max_L \pi = R(L) - w(L)L$$

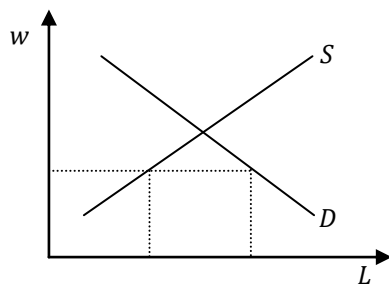
$$\Rightarrow \pi_L = R_L - \frac{\partial w}{\partial L} L - w = 0$$

$$\Rightarrow R_L = w + \frac{\partial w}{\partial L} L$$

$$\begin{aligned} \Rightarrow R_L &= w + \frac{1}{\frac{\partial L}{\partial w} L} \\ \Rightarrow R_L \frac{1}{w} &= 1 + \frac{1}{\frac{\partial L}{\partial w} L} \\ \Rightarrow R_L \frac{1}{w} &= 1 + \frac{1}{\varepsilon_{Lw}} \\ \Rightarrow R_L &= w \left(1 + \frac{1}{\varepsilon_{Lw}} \right) \end{aligned}$$

Hvis $\varepsilon_{Lw} < \infty$ ser vi at tilpasningen blir annerledes enn ved perfekt konkurranse der bedriften tar lønnsnivået som gitt. I så fall er $R_L > w$; det vil si at bedriften ansetter færre enn ved frikonkurranse ($R_L = w$). I så fall har bedriften monopsonimakt.

- Mål: Å estimere ε_{Lw} .
- Estimeringsproblem: En bedrift med monopsonimakt determinerer ε_{Lw} og w simultant.
- Identifikasjonsstrategier:
 - Å finne et instrument for L i relasjonen $w = \alpha_0 + \alpha_1 L + \alpha_2 x + u$. Problemet er imidlertid at det er svært vanskelig å finne instrumenter på bedriftsnivå som på den ene siden korrelerer med sysselsetting, men på den andre siden ikke korrelerer med lønn.
 - Å identifisere tilbud og etterspørsel under markedskrysset:



- Falch velger sistnevnte ved å se på læreransettelser i Nord-Norge.
 - Ansettelseslov: Kun lov å ansette lærere uten godkjent lærerutdanning dersom ingen lærere med slik utdanning søker (lærermangel). Dessuten kun lov å ansette slike ukvalifiserte lærere på ettårs-kontrakter.
 - Lønnslov: Lærerlønningene forhandles sentralt, derfor lik lønn for alle lærere. Imidlertid unntak for visse kommuner i Nord-Norge som ble gitt tillatelse til å øke lønnen hvis lærermangelen oversteg et visst nivå (ca. 20-30% av lærerstaben uten kvalifisert lærerutdanning).
 - Ansettelsesloven og lønnsloven gjør oss i stand til å identifisere både tilbudet og etterspørselen i skoler med ufaglærte lærere (etterspørselsoverskudd): Til gitt lønn vil tilbudet være gitt av antall lærere med kvalifisert lærerutdanning. Etterspørselen vil være gitt av antall lærere med totalt. Etterspørselsoverskuddet vil slik sett være differansen mellom totalt antall lærere og lærere med godkjent lærerutdanning, det vil si antall ufaglærte lærere.

- Klassifisering av skolestatus på et gitt tidspunkt (skoleår):
 - IN: Skoler som tilfredsstilte kravet til lønnspremie.
 - SIN: Skoler som har lærere med lønnspremie fordi de tidligere har vært IN, men som ikke gir lønnspremie til nye lærere fordi de ikke lenger kan karakteriseres som IN.
 - NIN: Skoler uten lærere med lønnspremie, som heller ikke tilfredsstiller kravet til slik premie ved nyansettelser.
- Skoler i behandlingsgruppe: De skolene som mottok lønnspremie for minst én av lærerne sine i løpet av måleperioden 1993-1994 til 2000-2001.
- Modellering:
 - Utrykk for den sanne tilbudskurven S_{it} :

$$S_{it} = (1 - q_{it})N_{it-1} + H_{it} + E_{it}^S = N_{it} + E_{it}^S = D_{it} - E_{it}^D$$
 der q_{it} er oppsigelsesraten, N_{it-1} er fjorårets sysselsetting, H_{it} er årets nyansetting, E_{it}^S er overskuddstilbudet, D_{it} er etterspørselen, E_{it}^D er etterspørselsoverskuddet. Så lenge $E_{it}^S = 0$ kan S_{it} identifiseres, siden vi har tall for $(1 - q_{it})N_{it-1} + H_{it} = N_{it}$.
 - Antar at q_{it} og H_{it} avhenger av lønnsnivået W_{it} . Dermed har vi følgende grunnleggende relasjon på bakgrunn av tilbudsuttrykket over (små bokstaver betyr at de er på logaritmeform):

$$s_{it} = \delta_i + \delta_t + \alpha n_{it-1} + \beta w_{it} + \gamma x_{it} + \eta_{it}^S \quad (1)$$
 der δ_i er skolespesifikke tidsuavhengige effekter, δ_t er skoleuavhengige tidsspesifikke effekter, w_{it} er lønnsnivået, x_{it} er kontrollvariabler og η_{it}^S er restleddet.
 - Merk: Lønnspremien betales av staten og har derfor ingen innvirkning på etterspørselen til skolene.
 - Problem med (1): Uobserverte forklaringsvariabler. Løser dette ved å estimere (1) ved hjelp av FE og FD.
- Resultater:
 - Deskriptiv statistikk:
 - Gjennomsnittlig lønnsvekst per lærer for IN- og SIN-skoler: 0,098
 - Fixed effects: $\hat{\beta} = \frac{0,108}{(0,026)}$
 - First difference: $\hat{\beta} = \frac{0,142}{(0,027)}$
- Forståelse:
 - Fixed effects: $\varepsilon_{Lw} = \frac{\partial L}{\partial w} \frac{w}{L} = 0,108 \times 0,098 \times 100 = 1,06$
 - First Difference: $\varepsilon_{Lw} = \frac{\partial L}{\partial w} \frac{w}{L} = 0,142 \times 0,098 \times 100 = 1,39$
- Konklusjoner:
 - Signifikant elasticitet som indikerer monopsonimakt.
 - Kritikkk: Kun kortsiktig elasticitet, og til en viss grad er dette kun skift i tilbud mellom skoler.

4.3. Johansen, Mydland & Strøm: Politics in wage setting: Does government color matter?

- Problemstilling: Opptreer fagforeninger mykere under lønnsoppgjørene når sosialdemokratiske konstellasjoner sitter med regjeringmakten?
- Teori:
 - Utgangspunkt: Sentraliserte oppgjør gir teoretisk sett mer moderat lønnsvekst enn desentraliserte oppgjør, og dermed også bedre oppgjør hensyn til sysselsetting, inflasjon og andre økonomiske variabler. Skyldes at nasjonale fagforeninger internaliserer makroøkonomiske konsekvenser av lønnsvekst.
 - I Skandinavia: Sentraliserte oppgjør med sterke bånd mellom fagforeninger og sosialdemokratiske partier (spesielt Arbeiderpartiet?). Antas at fagforeninger og sosialdemokratiske partier har lignende ideologisk plattform, der velferdsstaten og fellesskapet står sentralt.
 - Hypotese: Hvis regjeringen er sosialdemokratisk modererer fagforeningen lønnskravene i forventning om at økt økonomisk handlingsrom skal resultere i andre velferdstiltak. Motsatt; på grunn av manglende tillit til borgerlige regjeringer vil fagforeningene stille harde lønnskrav når disse styrer.
 - Konsekvens: Mer ekspansive lønnsoppgjør når borgerlige regjeringer styrer sammenlignet med sosialdemokratiske regjeringer.
 - Analytisk illustrasjon av Nash' forhandlingsløsning:

$$\Omega = (V(W, Z(W), Q) - \bar{V})^\lambda (\pi(W, X) - \bar{\pi})^{1-\lambda}$$
 der $V(W, Z(W), Q)$ er fagforeningens nytte som funksjon av lønnen W , sannsynligheten $Z(W)$ for en sosialdemokratisk regjering i kontraktperioden, og andre faktorer Q . Bedriftens profitt er en funksjon av lønn og andre faktorer X , og den relative forhandlingsmakten til fagforeningen er λ .
- Modellering:
 - Utgangspunkt i kvartalsdata:

$$\Delta_4 w_t = \Delta_4 \mathbf{x}\boldsymbol{\beta} - \alpha(w - p - pr)_{t-4} + \gamma(cpi - p)_{t-4} + f(\mathbf{U}) + \delta STOP_t + \lambda SOC_t$$
 - "Parameter of finterest": λ
 - Bruker tidsserie i stedet for panel for å unngå problemer knyttet til utelatte variable ved tverrsnitt, og siden det er vanskelig å definere sosialdemokratisk og borgerlig regjering generelt når man har med flere land. Dessuten antar paneldata at den lineære sammenhengen er lik på tvers av land. Dette er en urimelig antagelse i mange tilfeller, inkludert dette tilfellet.
 - Samspill: Tester også om avvik i arbeidsledigheten gir annen lønnsvekst under sosialdemokratisk styring.
- Resultater:
 - Effekten av SOC : $-\frac{0,010}{(0,002)}$
 - Effekten av $SOC_t(u_{t-3} - \bar{u})$: $-\frac{0,018}{(0,0,004)}$
- Forståelse:

- Effekten av en sosialdemokratisk regjering er 1% lavere lønnsvekst.
- Evnen til å justere seg tilbake etter et avvik fra likevektsledigheten er nesten 2% bedre når sosialdemokratiske regjeringer sitter med makten.

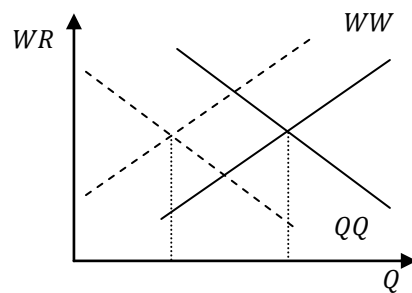
4.4. Johansen, Falch & Strøm: Teacher shortages and the business cycle

- Problemstilling:
 - Overordnet: Taper offentlig sektor i konkurransen om høykompetent humankapital?
 - Her: Sammenhengen mellom lærermangel og konjunktursykluser.
 - Tidligere studier tyder på at både den relative kvaliteten og relative lønninger har falt i offentlig sektor de siste tiårene.
- Teoretisk rammeverk:
 - Ved vekst i arbeidsmarkedet generelt kan det være vanskelig å holde på kvalifiserte arbeidstakere i offentlig sektor.
 - Problemet forsterkes av rigide offentlige lønnsordninger:
 - Ufleksibelt offentlig lønnsystem: Den relative offentlige lønnen faller i gode tider og stiger i dårlige tider.
 - Sentraliserte oppgjør: Homogent lønnsnivå skaper tilbudsoverskudd i noen regioner og etterspørselsoverskudd i andre.
 - Konsekvenser: Mangel på utdannet arbeidskraft i offentlig sektor i gode tider og på steder med "gode tider". Dermed variasjon i kvaliteten på det offentlige tjenestetilbudet!
 - Trenger altså data med variasjon i tid og geografi – paneldata.
- Mål på lærermangel: Skoler kan kun ansette lærere uten godkjent lærerutdanning hvis det ikke er noen søkere med slik utdanning. Derfor; $Q \equiv \frac{\# \text{ ikke-godkjente lærere}}{\# \text{ lærere totalt}}$.
- Modell:
 - Lønnsrelasjonen:

$$WR = a_1Q + a_1U + a_3X$$
 der $WR \equiv \frac{W^{teac\ her}}{W^{private}}$ er den relative lærerlønnen
 - $dQ > 0$ gir $dWR > 0$
 - $dU > 0$ gir $dWR > 0$ (fordi $W^{private}$ er mer følsom for konjunkturer enn $W^{teac\ her}$)
 - Lærermangelrelasjonen:

$$Q = b_1WR + b_2U + b_3D + b_4S$$
 der $Q \equiv \frac{\# \text{ ikke-godkjente lærere}}{\# \text{ lærere totalt}}$ er lærermangelraten
 - $dWR > 0$ gir $dQ < 0$
 - $dU > 0$ gir $dQ < 0$ fordi:
 - (1) Virker direkte gjennom b_2 .
 - (2) Virker indirekte via b_1 .

- Merk: Her har vi simultanitet.
- Grafisk:



- Estimeringsstrategier:
 - Tidsserietilnærming: Må bruke IV siden WR er endogen i (2).
 - Regionale paneldata: Utnytter at lønnsfastsettelsen ble sentralisert i perioden 1973-2002. Kan dermed kontrollere fullstendig for lærerlønninger ved hjelp av tidsdummyer.
- Tidsserietilnærming:
 - Strukturlikning:

$$Q_t = \alpha Q_{t-1} + \beta_1 \ln WR_t + \beta_2 \ln WR_{t-1} + \beta_3 \ln U_t + \beta_4 U_{t-1} + \beta_5 \ln \Delta \text{Stud}_t + \beta_0 + v_t$$
 - Problem: Simultanitet mellom Q og WR .
 - Løsning: IV
 - Instrumenter:
 - Førsteårskontraktsdummy: DUM_t
 - Endring i arbeidskraftproduktiviteten i privat sektor: $\Delta \ln ZS_t$
 - Lønnsfastsettelsen er muligens mer treg i offentlig sektor: $\ln U_{t-2}$
 - Redusert form:

$$\ln WR_t = d Q_{t-1} + d_1 \ln WR_t + d_2 \ln WR_{t-1} + d_3 \ln U_t + d_4 U_{t-1} + d_5 \ln \Delta \text{Stud}_t + d_6 DUM_t + \beta_7 \Delta \ln ZS_t + \beta_8 \ln U_{t-2} + d_0 + \varepsilon_t$$
 - Sargan-testen:
 - Beregner alle \hat{v}_t fra strukturlikningen der instrumentene inkluderes:

$$\hat{v}_t = a Q_{t-1} + a_1 \ln WR_t + a_2 \ln WR_{t-1} + a_3 \ln U_t + a_4 U_{t-1} + a_5 \ln \Delta \text{Stud}_t + a_6 DUM_t + a_7 \Delta \ln ZS_t + a_8 \ln U_{t-2} + a_0 + \text{error}_t$$
 - Finner $TR^2 \sim \chi_q^2$ der q er antall instrumenter (3).
 - Finner en kji-fordeling på $7,35 > 5,99$, altså forkastes H_0 om at alle instrumentene er ukorrelerte med v_t .
 - Forkaster DUM_t som instrument og finner da at Sargan-testen er ok.
 - Resultater:
 - OLS: $\hat{\beta}_3 = -1,645$
 - IV: $\hat{\beta}_3 = -1,889$

- Konklusjon: Sterk negativ effekt av ledighet på lærermangelen!
- Langsiktig effekt av en varig økning i arbeidsledigheten:
 - Lærermangel: $Q_t = Q_{t-1} = \bar{Q}$
 - Relativt lønnsnivå: $WR_t = WR_{t-1} = \overline{WR}$
 - Arbeidsledighet: $U_t = U_{t-1} = \bar{U}$
 - Strukturlikningen:

$$\bar{Q} = \alpha\bar{Q} + \beta_1 \ln\overline{WR} + \beta_2 \ln\overline{WR} + \beta_3 \ln\bar{U} + \beta_4 \bar{U} + \beta_5 \ln\Delta\text{Stud}_t + \beta_0 + v_t$$

$$\Rightarrow (1 - \alpha)\bar{Q} = (\beta_1 + \beta_2) \ln\overline{WR} + (\beta_3 + \beta_4) \ln\bar{U} + \beta_5 \ln\Delta\text{Stud}_t$$

$$\Rightarrow \bar{Q} = \frac{\beta_1 + \beta_2}{1 - \alpha} \ln\overline{WR} + \frac{\beta_3 + \beta_4}{1 - \alpha} \ln\bar{U} + \frac{\beta_5}{1 - \alpha} \ln\Delta\text{Stud}_t$$
 - Langsiktig effekt av en varig ledighetsøkning er dermed:

$$\frac{d\bar{Q}}{d\ln\bar{U}} = \frac{\hat{\beta}_3 + \hat{\beta}_4}{1 - \hat{\alpha}} = \frac{-1,889 + 1,240}{1 - 0,774} = -2,872$$
 - Forståelse: En økning i logaritmen til langsiktig ledighet på 1% gir en reduksjon i lærermangel på 2,8 %-poeng.
- Mulige utfordringer knyttet til tidsserie i denne forskningsrapporten:
 - Få observasjoner
 - Potensielle utelatte variable
 - Problemer med instrument (noe svak Sargan-test)
- Regional paneldatatilnærming:
 - Kan adressere følgende tilleggsspørsmål: Fører sentraliserte lærerlønninger til regional ubalanse av lærertilbud og -etterspørsel?
 - Strukturlikning:

$$Q_{it} = \beta_1 \ln\bar{W}_{jt} + \beta_2 \ln U_{jt} + \beta_3 D_{it} + \beta_4 S_{it} + \beta_5 t_j + \eta_i + \delta_t + v_{it}$$
 der Q_{it} er lærermangelen i kommune i på tidspunkt t . \bar{W}_{jt} er alternativlønnen i markedsregion j på tidspunkt t .
 - Bruker proxy for \bar{W}_{jt} : Industrilønnen i tiden 1985-1995, samt huspriser i tiden 1991-2002.
 - Resultater:
 - Pooled OLS: $\hat{\beta}_2 = \frac{0,144}{(0,556)}$
 - FE: $\hat{\beta}_2 = -\frac{0,946}{(0,278)}$
 - Forståelse:

$$dQ = \beta_2 \frac{1}{u} dU$$

$$\Rightarrow \frac{dQ}{dU} = \beta_2 \frac{1}{u}$$

$$\Rightarrow \frac{dQ}{dU} = \hat{\beta}_2 \frac{1}{\bar{u}} = -0,946 \frac{1}{2,6} = -0,364$$
 En økning på ett prosentpoeng i regional arbeidsledighet reduserer lærermangelen med ca. 0,35 prosentpoeng.
- Konklusjon: Variasjonen i lærermangelen kan kobles til variasjon i ledighet.

5. Education Economics

5.1. Angrist & Lavy: Using Maimonides' rule to estimate the effect of class size on scholastic achievement

- Problemstilling: Effekten av klassestørrelse på elevprestasjoner.
- Hypotese: I små klasser er det mer lærertid på hver elev. Dette gir bedre elevprestasjoner.
- Modell: $y_{sc} = X'_s\beta + \alpha n_{sc} + \eta_s + u_{sc}$
der y_{sc} er elevprestasjoner i klasse c på skole s , X'_s er en vektor for kontrollvariabler, n_{sc} er klassestørrelsen, η_s er skolespesifikke egenskaper og u_{sc} er restleddet.
- Estimeringsproblem: Endogenitet knyttet til klassestørrelsen:
 - Utgangspunkt: Antar at effekten av klassestørrelse på elevprestasjoner er positiv.
 - Utelatt variabel: Læringsproblemer
 - Elever med læringsproblemer kan puttes systematisk i mindre klasser, altså negativ korrelasjon mellom klassestørrelse og læringsproblemer.
 - I så fall får vi oppadgående skjevhet i den estimerte effekten.
 - Utelatt variabel: Lokale sosiale problemer
 - Områder med sosiale problemer kan ha større klasser, altså positiv korrelasjon mellom klassestørrelse og sosiale problemer.
 - I så fall får vi nedadgående skjevhet i den estimerte effekten.
- Identifikasjonsstrategi: Å instrumentere klassestørrelsen ved hjelp av Maimonides klassestørrelsesregel.
- Regelen:
 - Utgangspunkt: Maksimalt 40 studenter per klasse i Israel.
 - Funksjon: $f_{sc} = \frac{e_s}{\text{int}\left(\frac{e_s-1}{40}+1\right)}$
der f_{sc} er estimert klassestørrelse i klasse c i skole s , e_s er elevopptak på skolen, og $\text{int}\left(\frac{e_s-1}{40}+1\right)$ er nærmeste heltall som er $\leq \left(\frac{e_s-1}{40}+1\right)$.
 - Eksempler:
 - 39 elever: $f_{sc} = \frac{39}{\text{int}\left(\frac{39-1}{40}+1\right)} = \frac{39}{1} = 39$
 - 40 elever: $f_{sc} = \frac{40}{\text{int}\left(\frac{40-1}{40}+1\right)} = \frac{40}{1} = 40$
 - 41 elever: $f_{sc} = \frac{41}{\text{int}\left(\frac{41-1}{40}+1\right)} = \frac{41}{2} = 20,5$
- Egenskap i israelske skoler: Positiv sammenheng mellom e_s og elevprestasjoner. Skyldes antageligvis at store skoler ligger i urbane og velstående områder.
- Resultater med OLS: Sterk positiv korrelasjon mellom klassestørrelse og elevprestasjoner, men denne blir insignifikant når vi kontrollerer for PD (percent disadvantaged).
- IV-estimering:
 - Første steg: $\hat{n}_{sc} = \pi_1 X'_s + \pi_2 f_{sc} + \xi_{sc}$
 - Andre steg: $y_{sc} = X'_s\beta + \alpha \hat{n}_{sc} + \eta_s + u_{sc}$

- Resultater: Negativ og signifikant sammenheng.
- Robusthetstest: Diskontinuitetstilnærming.
 - Bruker kun observasjonene rundt diskontinuitetspunktene $e_s = 40, 80, 120$.
 - Fortsatt negativ og signifikant sammenheng.
- Konklusjon: Finner en substansiell negativ effekt av klassestørrelsen på elevprestasjoner.

5.2. Black, McKinnish & Sanders: Tight labor markets and the demand for education

- Problemstilling: Hvordan avhenger etterspørselen etter utdanning av arbeidsmarkedsmuligheter (alternativavkastningen på utdanning)?
- Teori:
 - Utdanning kan tolkes som et investeringsprosjekt; utdanning er en investering i egen humankapital.
 - Elever må velge mellom å fullføre videregående eller falle fra videregående. Hvis avkastningen av videregående er størst velges fullføring av videregående.
 - Inntekt ved å fullføre: $E_{jg} = \sum_{t=k}^T \frac{w_{gt}}{(1-\rho_j)^t}$
 - Inntekt ved å falle fra: $E_{jd} = \sum_{t=0}^T \frac{w_{dt}}{(1-\rho_j)^t}$
 - Anta videregående utdanning fra år $t = 0$ til og med $t = k - 1$. Fra og med år $t = k$ er det jobb uansett. Elev j fullfører videregående så lenge den forventede livstidsinntekten av dette valget er størst, altså så lenge:

$$\sum_{t=k}^T \frac{w_{gt}}{(1-\rho_j)^t} - \sum_{t=0}^T \frac{w_{dt}}{(1-\rho_j)^t} > 0$$
 der

$$\sum_{t=k}^T \frac{w_{gt} - w_{dt}}{(1-\rho_j)^t}$$
 er avkastningen av en videregående grad og

$$\sum_{t=0}^{k-1} \frac{w_{dt}}{(1-\rho_j)^t}$$
 er tapt inntekt som følge av videregående:

$$\sum_{t=k}^T \frac{w_{gt} - w_{dt}}{(1-\rho_j)^t} - \sum_{t=0}^{k-1} \frac{w_{dt}}{(1-\rho_j)^t} > 0$$
 - Effekten av lønnsjokk:
 - w_g og w_d opp like mye:
 - Ser at avkastningen på utdanning er uendret, men at alternativkostnaden øker.
 - Sannsynligvis liten effekt bortsett fra når ρ er høy.
 - Midlertidig økning i w_d :
 - Midlertidige effekter har gjerne lite å si i et livstidsperspektiv.
 - Permanent endring i $w_g - w_d$:
 - Vil gi proporsjonale endringer i avkastningen knyttet til utdanning.
 - Kan få store konsekvenser for opptak og frafall i videregående.
- Modellering:

- Enkel relasjon: $ENROL_{ct} = b_0 + b_1 w_{dt} + b_2 w_{gt} + b_3 andre_{ct} + u_{ct}$
- Forventning: $b_1 < 0$ og $b_2 > 0$
- Mulig restriksjon: $b_1 = -b_2$
- Problemer:
 - Antageligvis er $ENROL_{ct}$ påvirket av utelatte variable, som også kan korrelere med lønn.
 - Siden teorien foreslår at permanente lønnsforskjeller må til for å gi effekt, nytter det ikke å bruke FD.
- Løsning: Å bruke konjunkturer i kullmarkedet som instrument for lønnsforskjellene.
- Deskriptivt om kullprisene:
 - Relativt stabile frem til 1969.
 - Stor variasjon etter 1969 grunnet nye reguleringer og grunnet OPECs oljeprissjokk.
- Modellering:
 - Sammenligner 1970 og 1980.
 - Redusert form-likning:

$Wage_i$

$$\begin{aligned}
 &= \sum_{j=10}^{12} \beta_{1j} d(j)_i + \sum_{j=10}^{12} \beta_{2j} d(j)_i DC_i + \sum_{j=10}^{12} \beta_{3j} d(j)_i D80_i \\
 &+ \sum_{j=10}^{12} \beta_{4j} d(j)_i DC_i D80_i + \beta_{5j} \mathbf{x} + u_i
 \end{aligned}$$

der $d(j)_i$ er en indikator for antall fullførte skoleår, DC_i er en dummy for kullområder, $D80_i$ er en dummy for året 1980, og \mathbf{x} er en vektor for kontrollvariabler.

- Kan tolke β_{4j} som en difference-in-difference-estimator:
 - Lønn utenfor kullområder i 1970: β_{1j}
 - Forskjell i lønn mellom områder med og uten kull i 1970: β_{2j}
 - Endring over tid for alle: β_{3j}
 - Effekten av kullområde (behandling) under boomen: β_{4j}
 - Finner at β_{4j} er langt større for $j = 10$ enn $j = 12$, det vil si for de som falt fra videregående.
- Så langt: Forventer å se relativ nedgang på videregående skoleopptak i kulldriftsområder under kull-”boom” sammenlignet med områder med lite kulldrift.
- Modellering:

- Utgangspunkt:

$$FrHS_{cst} = \beta_0 + \beta_c + \beta_1 EARNPW_{cst} + \beta_2 FrGS_{cst} + \beta_3 (State_s Year_t) + \varepsilon_{cst}$$

- Tar FD for å kontrollere for individspesifikke tidsuavhengige faktorer:

$$\Delta FrHS_{cst} = \beta_1 \Delta EARNPW_{cst} + \beta_2 \Delta FrGS_{cst} + \beta_3 (State_s Year_t) + \Delta \varepsilon_{cst}$$

- Problem: $\Delta EARNPW_{cst}$ vil i stor grad reflektere kortsiktige fluktuasjoner i lønn. Løsningen blir en IV-tilnærming.

- IV-tilnærming:

- Strukturlikningen:

$$\Delta FrHS_{cst} = \beta_1 \Delta \widehat{EARNPW}_{cst} + \beta_2 \Delta FrGS_{cst} + \beta_3 (State_s Year_t) + \Delta \varepsilon_{cst}$$

- Instrument: Kullverdien

- Redusert form:

$$\Delta EARNPW_{cst} = \alpha_1 \Delta COALVALUE_{cst} + \alpha_2 \Delta FrGS_{cst} + \alpha_3 (State_s Year_t) + \Delta \eta_{cst}$$

- To varianter av instrumentet:

- Kullreserver og realprisen på kull:

$$\begin{aligned} EARNPW_{cst} &= \alpha_1 \ln COALRESERVES_{cst} \Delta \ln (REALPRICE_{cst} \\ &\quad - REALPRICE_{cst-1}) + \alpha_2 \Delta FrGS_{cst} + \alpha_3 (State_s Year_t) \\ &\quad + \Delta \eta_{cst} \end{aligned}$$

- Interaksjon mellom 3 dummyer for kullområde og 3 dummyer for tidsperioder:

$$\begin{aligned} &D_{BIG}, D_{MED}, D_{SMA} \text{ og} \\ &D_{70-77}, D_{28-82}, D_{83-90} \text{ (boom, peak, burst)} \end{aligned}$$

- Resultater:

- FD: $\hat{\beta}_1 = \frac{0,003}{(0,034)}$

- IV-1: $\hat{\beta}_1 = -\frac{0,648}{(0,179)}$ med F -test på 22,4 i førstesteget (godt instrument)

- IV-2: $\hat{\beta}_1 = -\frac{0,490}{(0,104)}$ med F -test på 13,3 i førstesteget

- Forståelse: En økning i inntekt per arbeider $EARNPW$ på 1 %-poeng gir en reduksjon i videregåendeopptaket på 0,5-0,65 %-poeng. Merk: En økning i $EARNPW$ på ett prosentpoeng kan representere en økning i w_d på mer enn ett prosentpoeng.

- Konklusjoner:
 - Langsiktige variasjoner i relativ inntekt driver etterspørselen etter utdanning.
 - Policy-implikasjon: Minimumslønn og andre sosialøkonomiske tiltak kan ha den utilsiktede effekten at investeringen i humankapital svekkes.
- Sammenlignet med Norge:
 - I Norge har vi langt mer rigide lønninger enn i USA. Dermed gir endringer i for eksempel fiskepriser større effekter på ledigheten og mindre effekter på lønn enn i USA.
 - Tilhørende kritikk: Skulle tatt med arbeidsledighetsrate i modellen.

5.3. Leuven, Oosterbeek & Rønning: Quasi-experimental estimates of the effect of class size on achievement in Norway

- Problemstilling: Effekten av klassestørrelse på elevprestasjoner.
- Forskjell fra Angrist & Levy:
 - Klassestørrelsesregelen
 - Populasjonsvariasjon
 - Mindre ungdomsskoler
 - Kjent klassestørrelse over tre år
 - Norske data
- Motivasjon for en separat norsk analyse: Resultater viser seg å variere mellom land og mellom studier.
- Strukturmodell:

$$y_i = x_i' \beta + w_{s(i)}' \alpha + \delta \bar{c}_{s(i)} + \eta_{s(i)} + \psi_{t(i)} + \varepsilon_i$$

Der y_i er elevprestasjon for elev i , x_i' er en vektor for elevspesifikke karakteristika, $w_{s(i)}'$ er en vektor for skolespesifikke karakteristika, $\bar{c}_{s(i)}$ er gjennomsnittlig klassestørrelse, $\eta_{s(i)}$ er uforklart skoleeffekt, $\psi_{t(i)}$ er åreffekt og ε_i er restleddet.
- Problem: Endogenitet i klassestørrelse.
- Identifikasjon 1: Regel for klassestørrelse som instrument for faktisk klassestørrelse
 - Regel: $f_{sc} = \frac{e_s}{\text{int}\left(\frac{e_s-1}{30}+1\right)}$
 - Redusert form:

$$\bar{c}_{s(i)} = x_i' \pi_1 + w_{s(i)}' \pi_2 + \delta f_{s(i)} + \text{error}_i$$
- Identifikasjon 2: Utnytter populasjonsvariasjon
 - Antar et gitt # klasser og tenker oss at klassestørrelsen varierer med variasjonen i opptak.
 - Opptak:
 - Deterministisk: Folk flytter, osv.

- Tilfeldig: Tidspunkt for fødsler, osv.
 - Idé: Å bruke tilfeldig del av opptak som instrument.
 - Formelt:

$$e_{st} = \bar{e}_{st} u_{st}$$
 der e_{st} er opptak, \bar{e}_{st} er den deterministiske delen og u_{st} er den tilfeldige delen.

$$\Rightarrow \ln e_{st} = \ln \bar{e}_{st} + \ln u_{st}$$
 Der estimatet til $\ln u_{st}$ kan brukes som instrument for klassestørrelse.
 - Hjelperegresjon:

$$\widehat{\ln u}_{st} = \ln e_{st} - \sum_{k=0}^K \alpha_{sk} t^k$$
 - Førstesteget:

$$\bar{c}_{s(i)} = x'_i \pi_1 + w'_{s(i)} \pi_2 + \delta \widehat{\ln u}_{st} + \text{error}_i$$
 - Andresteget:

$$y_i = x'_i \beta + w'_{s(i)} \alpha + \delta \bar{c}_{s(i)} + \eta_{s(i)} + \psi_{t(i)} + \varepsilon_i$$
- Andre strategier:
 - Analyser på skoledistriksnivå for å kontrollere for foreldres strategiske flytting av barn mellom skoler.
 - Å kun bruke små skoler (med under 20 elever per trinn) der klassestørrelsen varierer direkte med e_s .
 - Kombinasjon av klassestørrelsesregelen og tilfeldig variasjon i antall studenter.
 - Bruker lærertime per student i stedet for klassestørrelse.
- Resultater:
 - Generelt: Ingen effekt av klassestørrelse på elevprestasjoner.
 - Tester for heterogene effekter av klassestørrelse ved å splitte utvalg etter:
 - Kjønn
 - Foreldres utdanning
 - Minoritetsstatus
 - Lærerutdanning
 og finner lite som tyder på heterogene effekter.
 - Overordnet: Lite tyder på klassestørrelseeffekt hos norske tiendeklassinger.
- Hvis det likevel kan være klassestørrelseeffekter kan dette skyldes en av følgende to:
 - Produksjonsteknologi:
 - Effekten av klassestørrelse når andre innsatsfaktorer holdes konstant.
 - Hvis det er slik at lærere ikke evner å utnytte mindre klasser kan dette utbedres gjennom lærerutdanning.
 - Policy:
 - Effekten av klassestørrelse når andre innsatsfaktorer ikke holdes konstant.
 - Hvis foreldre substituerer redusert lærerinnsats med økt foreldreinnsats er det ikke mye å hente på å redusere klassene.

5.4. Monstad, Propper & Salvanes: Education and fertility: Evidence from a natural experiment

- Problemstilling: Påvirker økt utdanningsnivå kvinners fertilitet?
- Bakgrunn: Debatt i mediene om at kvinners økende utdanningsnivå har ført til lavere fødselsrate.
- Deskriptiv observasjon: Utdanning og fertilitet er negativt korrelert.
- Identifikasjonsstrategi: Å bruke utdanningsreform i Norge for å estimere den kausale effekten av utdanning på fertilitet.
- Reform: Reformen i 1960-1970-årene som forlenget obligatorisk skolegang.
- Fertilitetsmål:
 - Tidspunkt for fødsler
 - Barnløshet
 - Antall barn
- Teori:
 - Avgjørelsen om barn inngår i en standard nyttefunksjon: $u(\#barn, barnekvalitet, andre\ goder)$ under budsjettbetingelser
 - Kostnads-nyttesammenligning:
 - Høyere utdanning
 - ⇒ Høyere alternativkostnad ved å oppdra barn
 - ⇒ # barn ned, barnekvalitet opp
 - Empirisk spørsmål: Gevinster og kostnader ved tidlig fødsel.
 - Mulig: Økt utdanning gir mer å tape på å føde tidlig.
- Strukturmodell:

$$y_i = \beta_0 + \beta_1 ED_i + \beta_2 COHORT_i + \beta_3 MUN_j + \beta_4 TREND_i + \varepsilon_{ij}$$
 - OLS på denne modellen direkte krever at $Cov(\varepsilon_{ij}, ED_i) = 0$. Dette er lite sannsynlig grunnet utelatte variabler, for eksempel "evne". Kvinner med evne investerer både mer i utdanning og foretrekker færre barn.
 - Vanskelig å observere evne.
 - Løsning: IV
- Bruker reform som IV.
- Redusert form:

$$ED_i = \alpha_0 + \alpha_1 REFORM_j + \alpha_2 COHORT_i + \alpha_3 MUN_j + \alpha_4 TREND_i + \eta_{ij}$$
 - Så lenge α_1 er signifikant er instrumentet relevant, men det må også være eksogent: $Cov(\varepsilon_{ij}, REFORM_i | COHORT_i, MUN_j, TREND_i) = 0$.
 - Mulig problem i forhold til eksogenitetskravet: Migrasjon! Også mulige målefeil; man folkeregistreres ofte senere enn man flytter fysisk.
- Resultater:
 - Ikke signifikant effekt verken på antall barn eller barnløshet.
 - Signifikant færre som føder sin førstefødte før fylte 20, og flere som føder etter fylte 35. Ser ut som utdanning bare fører til senere fødsler.