Økonometri HO3

Jeppe Vanderhaegen, wlr139 og Jacob Vestergaard, gcb906

9. maj 2022

Opgave 1

Vi tager udgangspunkt i følgende regressionsmodel:

$$\begin{aligned} dod70 = &\beta_0 + \beta_1 pension66 + \beta_2 \log(indkomst) + \beta_3 kvinde + \beta_4 faglaert \\ &\Box + \beta_5 kortudd + \beta_6 mludd + u \end{aligned}$$

(a) Regressionmodellen præsenteret i opgaveformuleirngen er en lineær sandsynlighedsmodel. Modellen ønsker at beskrive sandsynligheden for at dø, inden man fylder 70 år. Modellen består af flere forklarende variable, og er derfor en MLR model. De forklarende variable består af en dummy for køn, tre forskellige uddannelses dummy-variabler, som hver islr opsummerer niveauet og længden af forskellige udannelser præsenteret i datasættet, en dummy for om hvorvidt vedkommende er pensioneret før alderen 66 år, og til sidst det enkelte individs indkomst, som er blevet log-transformeret. Da *dod*70 er en dummy-variabel, vil β-estimaterne måle den forventede ændring i sandsynligheden for, at man dør inden man er 70 år gammel, som følge af at den pågældende forklarende variable er aktiv/til stede. Derfor er fortolkningen af β_1 og β_3 som følgende. β_1 angiver ændringen i sandsynligheden for at dø inden man fylder 70, som følge af at man er pensioneret før alderen 66. β_3 angiver ændringen i sandsynligheden for at dø inden man fylder 70, hvis man er en kvinde (i modsætningen til at være en mand).

(b) Vi konstruerer dummy-variablen (FP), som er 1 for personer hvor den officielle pensionsalder er 65 år, og 0 for de personer hvor den officielle pensionsalder er 67 år. Tilsvarende gøres dette for uddannelse. På baggrund af den nedenstående tabel, vil følgende deskriptive analyse af data foretages. Tabellen er opdelt efter observationerns officielle folkepensionsalder. Standardafvigelse og gennemsnit er rapporteret.

	pens. alder 67		pens. alder 65		Alle	
	mean	sd	mean	sd	mean	sd
dod70	0.133	0.340	0.133	0.340	0.133	0.340
pension66	0.513	0.500	0.513	0.500	0.513	0.500
indkomst	208569.486	156374.012	208569.486	156374.012	208569.486	156374.012
fodaar	1939.018	0.823	1939.018	0.823	1939.018	0.823
kvinde	0.510	0.500	0.510	0.500	0.510	0.500
faglaert	0.391	0.488	0.391	0.488	0.391	0.488
kortudd	0.137	0.344	0.137	0.344	0.137	0.344
mludd	0.033	0.178	0.033	0.178	0.033	0.178
No. obs	1530		1530		1530	

Table 1: Deskriptiv analyse af data

Ved første øjekast på tabellen, er der umiddelbart ikke den store forskel på de to grupperinger. Dog kan det ses, at omtrent 39 procent af dem med en folkepensionsalder på 67 år, pensioneres før de er 66 år gammel. For dem med en folkepensionsalder på 65, er dette resultat på 63 procent, hvilket umiddelbart er en betydelig forskel.

(c) Vi estimerer nu modellen med OLS i STATA. Da der også foretages regressioner i opgave 2, kan resultatet af denne regression observeres i den samlede tabel, som er fremført i næste opgave. Da den variabel vi ønsker at forklare, $dod70$, er en dummy-variabel, vil der per definition, altid være heteroskedasticitet i vores model, jf. at $dod70$ er en Bernoulli fordelt stokastisk variabel. Derfor foretages der en robust OLS estimation. Teoretisk set, så vil estimaterne af vores β-værdier angive den forventede ændring i sandsynligheden for at dø, inden man fylder 70 år. Estimaterne er konsistente, såfremt at MLR.1-4 er overholdt (og middelret for den sags skyld). Vi vil nok ikke forvente at estimaterne er konsistente, da der højst sandsynligt er omvendt kausalitet i vores model. På grund af dette, vil variansen af fejlleddet afhænge af vores forklarende variable, som bryder med MLR.4 antalgesen.

Opgave 2

(a) Modellen estimeres nu ved en ved en IV-estimation, hvor dummien for den officielle folkepensionsalder,FP, anvendes som instrument for pension66. FP kan bruges som instrument, hvis variablen er korreleret med variablen *pension66*. Herved opstiller vi en hjælperegression, hvor i der testes for hypoteserne:

- $\mathcal{H}_0 : FP = 0$
- $\mathcal{H}_A : FP \neq 0$

Parameteren FP estimeres til 0,247 med en t-værdi på 10,29 og p-værdi på 0. På et 5% signifikant-niveau forkastes nulhypotesen, da estimatet er signifikant på alle konventionelle signifikansniveauer, da p-værdien er 0. Dermed er FP korreleret med den endogene variabel pension66. Ydermere skal FP være ukorreleret med fejlleddet, før FP kan være et instrument. I og med vi kigger på en reformændring antager vi, at personerne inkluderet, ikke har haft nogen effekt på ændringen. Hermed mener vi, at FP kan være et instrument, da det opfylder kriterierne (korreleret med den endogene variabel og ukorreleret med fejlleddet)

(b) Vi opdeler nu variationen i potentiel endogene variabel i en eksogen del og en potentiel endogen del. Indgår den potetielt endogene del i modellen, vil variablen være endogen. Residualerne fra den tidligere regression er et estimat på den potentielt endogene del. Hermed regresseres igen hvori residualerne indgår, hvorved paramterværdier og standardafvigelse observeres. Vi opstiller hermed en nulhypotese, der siger *pension66* er en eksogen variabel. Dette testes ved en z-test.

$$z = \frac{\beta_i - \mathcal{H}_0}{se(\hat{\beta}_i)} = \frac{0,167 - 0}{0,074} = \frac{0,167}{0,074} = 2,26$$

Dette er større end den kritiske værdi på et 0,05 signifikans niveau. Herved forkastes nulhypotesen, og hermed er variablen *pension66* endogen. Af denne grund foretrækkes IV-estimation. IV-estimatet for $\beta_1 = -0,073$, og vha. af en Wald-test, kan det udledes, vi ikke kan forkaste $\beta_1 = 0$. Dette betyder, at tidlig pension ikke har en kausal effekt på en tidlig død. IV-estimatet for $\beta_3 = -0,068$. Vha. konfidensintervallet kan det ses, at parameteren er markant forskellig fra nul. Hermed kan det udledes, at kvinder har en $6,8\%$ mindre chance for at dø tidligere end mænd.

(c) Histogrammet kan observeres på sidste side i dokumentet, udtryk som figur 1. Histogrammet viser den prædikterede sandsynlighed for en tidlig død for hhv. mænd og kvinder. Det kan udledes, at der kun er enkelte observationer for kvinder, der ligger under nul. Observationerne er marginale ift. variationen, hvorved vi vurderer modellen til at være velspecificeret.

Opgave 3

(a) Vi benytter nu følgende model:

$$\begin{aligned} laege70 = &\gamma_0 + \gamma_1 pension66 + \gamma_2 \log(indkomst) + \gamma_3 kvinde + \gamma_4 faglaert \\ &\Box + \gamma_5 kortudd + \gamma_6 mludd + u \end{aligned}$$

Denne model estimeres ved en IV estimation, hvor dummien for den officielle folkepensionsalder, FP, anvendes som instrument for $pension66$. Resultaterne kan ses i den samlede tabel, hvor alle regressionerne forbundet med hele opgaven er samlet.

	dodOLS	dodIV	dod_{test}	laegeIV	laegeIV2
pension66	0.083***	-0.073	-0.073	-0.535***	-0.141
	(0.018)	(0.072)	(0.070)	(0.150)	(0.222)
logind	-0.038*	-0.033*	-0.033*	-0.022	-0.024
	(0.016)	(0.016)	(0.016)	(0.034)	(0.031)
kvinde	-0.085***	-0.068***	-0.068***	-0.388***	-0.035
	(0.019)	(0.020)	(0.020)	(0.051)	(0.148)
faglaert	-0.018	-0.045	-0.045*	-0.041	-0.040
	(0.020)	(0.023)	(0.023)	(0.055)	(0.052)
kortudd	-0.041	-0.082**	-0.082**	0.125	0.120
	(0.025)	(0.031)	(0.030)	(0.077)	(0.073)
mludd	0.016	-0.042	-0.042	0.317*	0.357**
	(0.054)	(0.060)	(0.058)	(0.132)	(0.120)
Residuals			0.167*		
			(0.074)		
kvindepen					-0.733*
					(0.297)
Constant	0.605**	0.636**	0.636**	1.421***	1.290***
	(0.193)	(0.198)	(0.193)	(0.407)	(0.385)
Obs	1530	1530	1530	1326	1326
Rsq	0.032	.	0.036	.	0.091

Table 2: Resultater af estimationer

Ovenover kan regression observeres (og alle andre der er fuldført). Det ses fra tabellen, at antallet af lægebesøg er aftagende, såfremt at vedkommende er pensioneret før 66 år. Antallet af lægebesøg falder i gennemsnittet med $-0,535$, dvs. et halvt lægebesøg pr år, som følge af at det enkelte individ i data er pensionereret før 66 år. Det kan ses fra tabellen, at vi har at gøre med 1329 observationer for $laege70$. Førhen havde vi at gøre med 1530, hvilket var betydeligt mere. Da vi i modellen har specifieret antallet af lægebesøg efter fyldt 70, vil det betyde, at individer som er døde inden de fyldte 70 år, er udgået fra data. Umiddelbart kan dette godt resultere i problemer med regressionen, da det naturligvis betyder at vi observerer et reduceret datasæt, hvor vi har elimineret de meget syge personer.

(b) For at teste om uddannelse har en effekt på antallet af lægebesøg opstilles følgende nulhypotese og alternativ hypotese:

- $\mathcal{H}_0 : \gamma_4 \wedge \gamma_5 \wedge \gamma_6 = 0$
- $\mathcal{H}_A : \gamma_4 \wedge \gamma_5 \wedge \gamma_6 \neq 0$

F-teststørrelsen bliver 12,99. Denne er asymptotisk $\mathcal{X}^2$-fordelt med tre frihedsgrader, og har en p-værdi på 0,0047. Da p-værdien er marginalt tæt på 0, vil vi på et 5% signifikant-niveau forkaste nulhypotesen, da estimatet er signifikant på alle konventionelle signifikansniveauer. Dermed må længden af uddannelse have en betydning for antallet af lægebesøg.

(c) Modellen er estimeret ved en IV-regression. Her er *FP* og *FP*kvinde* brugt som instrumenter for *pension66* og *pension66*kvinder*. Resultaterne er vist i tabel 2, kolonne 5.

Nu testes der for, om pensionering har forskellig effekt på mænd og kvinder. Herved opstiller vi følgende hypoteser:

- $\mathcal{H}_0 : \gamma_7 = 0$
- $\mathcal{H}_A : \gamma_7 \neq 0$

Det ses, at t-værdien er -2,47. Den kritiske værdi for en asymptotisk standard normalfordeling på et 5% signifikansniveau kendes som 1,96. Herved kan vi forkaste nulhypotesen. Dette betyder, der er forskel på effekten af pensionering mellem mænd og kvinder. Det kan udledes, at antallet af kvinders lægebesøg falder med -87 gange per år, mens antallet af mændenes falder med $0,14$ gange per år ved et 50% signifikansniveau, hvorved det ikke er signifikant.

Opgave 4

(a) Vi betragter følgende regressions model:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Vi antager at $E(u) = 0$ og $E(ux) = 0,10$. Vi antager at x er en dummy-variabel. For x gælder der, at $E(x) = P(x = 1) = 0,15$. Variansen beregnes som følgende:

$$Var(x) = E(x^2) - E(x)^2 \Leftrightarrow$$

$$Var(x) = \left(0,85 \cdot 0^2 + 0,15 \cdot 1^2\right) - 0,15^2 \Leftrightarrow$$

$$Var(x) = 0,15 - 0,0225 \Leftrightarrow$$

$$Var(x) = 0,1275$$

Hermed er variansen af x udledt. Den asymptotiske bias af OLS estimatoren for β_1 kan udledes ved følgende formel:

$$p\lim \hat{\beta}_1 - \beta_1 = \frac{Cov(x,u)}{Var(x)}$$

Vi kan dekomponere $Cov(x, u)$, samt indsætte værdier for de enkelte komponenter:

$$p\lim \hat{\beta}_1 - \beta_1 = \frac{E(ux) - E(x) \cdot E(u)}{Var(x)} \Leftrightarrow$$

$$p\lim \hat{\beta}_1 - \beta_1 = \frac{0,10 - 0,15 \cdot 0}{0,1275} \Leftrightarrow$$

$$p\lim \hat{\beta}_1 - \beta_1 = 0,7843$$

Hermed er den asymptotiske bias af OLS estimatoren for β_1 udledt.

(b) Da vi betragter en SLR model, ved vi at den generelle estimator for $\hat{\beta}_1$ er givet som:

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

Ved brugen af gængse hjælpesætninger for summer, kan dette udtryk skrives som (tjek noter):

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

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

Benytter hintet fra opgaveformuleringen:

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

Da x er en dummy-variabel, gælder der at $x_i^2 = x_i$, da både værdien af 0 og 1 er uændret, når det opløftes i 2. Derfor kan vi skrive udtrykket som:

$$\hat{\beta}_1 = \frac{n\bar{x}\bar{y}^{x=1} - \bar{x}\bar{y}n}{\sum_{i=1}^{n} x_i - \bar{x}\bar{x}n} \Leftrightarrow$$

$$\hat{\beta}_1 = \frac{n\bar{x}\bar{y}^{x=1} - \bar{x}\bar{y}n}{\bar{x}n - \bar{x}\bar{x}n}$$

Indsætter definitionen af $\bar{y}$, som er givet fra opgaveformuleringen:

$$\hat{\beta}_1 = \frac{n\bar{x}\bar{y}^{x=1} - \bar{x}n(\bar{x}\bar{y}^{x=1} + (1-\bar{x})\bar{y}^{x=0})}{\bar{x}n - \bar{x}\bar{x}n} \Leftrightarrow$$

$$\hat{\beta}_1 = \frac{n\bar{x}\bar{y}^{x=1} - \bar{x}n\bar{x}\bar{y}^{x=1} - \bar{x}n(1-\bar{x})\bar{y}^{x=0}}{\bar{x}n(1-\bar{x})} \Leftrightarrow$$

$$\hat{\beta}_1 = \frac{n\bar{x}(\bar{y}^{x=1} - \bar{x}\bar{y}^{x=1} - (1-\bar{x})\bar{y}^{x=0})}{\bar{x}n(1-\bar{x})} \Leftrightarrow$$

$$\hat{\beta}_1 = \frac{n\bar{x}(1-\bar{x})(\bar{y}^{x=1} - \bar{y}^{x=0})}{\bar{x}n(1-\bar{x})} \Leftrightarrow$$

$$\hat{\beta}_1 = \bar{y}^{x=1} - \bar{y}^{x=0}$$

Hermed er OLS estimatoren for $\hat{\beta}_1$ udledt, som er ækvivalent med opgave formuleringen.

(c) Vi skal vise at estimatoren for $\tilde{\beta}_1$ er konsistent estimator for β_1. Estimatoren er givet som:

$$\tilde{\beta}_1 = \frac{\bar{y}^{d=1} - \bar{y}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}}$$

Vi ved at $\bar{y}^{d=1}$ kan skrives som:

$$\bar{y}^{d=1} = \frac{\sum_{i=1}^{n} y_i d_i}{\sum_{i=1}^{n} d_i}$$

Da MLR.1 er overholdt, kan vi indsætte denne i ovenstående:

$$\bar{y}^{d=1} = \frac{\sum_{i=1}^{n} (\beta_0 + \beta_1 x_i + u_i) d_i}{\sum_{i=1}^{n} d_i} \Leftrightarrow$$

$$\bar{y}^{d=1} = \frac{\sum_{i=1}^{n} \beta_0 d_i + \beta_1 x_i d_i + u_i d_i}{\sum_{i=1}^{n} d_i} \Leftrightarrow$$

$$\bar{y}^{d=1} = \beta_0 \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} d_i} + \beta_1 \frac{\sum_{i=1}^{n} x_i d_i}{\sum_{i=1}^{n} d_i} + \frac{\sum_{i=1}^{n} u_i d_i}{\sum_{i=1}^{n} d_i}$$

Bruger definitionen af $\bar{x}^{d=1}$, som er givet fra opgaveformuleringen. Analogt bruges denne, for $\bar{u}^{d=1}$:

$$\bar{y}^{d=1} = \beta_0 + \beta_1 \bar{x}^{d=1} + \bar{u}^{d=1}$$

Samt gælder der for $\bar{y}^{d=0}$:

$$\bar{y}^{d=0} = \beta_0 + \beta_1 \bar{x}^{d=0} + \bar{u}^{d=0}$$

Disse to definitioner indsættes i $\tilde{\beta}_1$:

$$\tilde{\beta}_1 = \frac{\beta_0 + \beta_1 \bar{x}^{d=1} + \bar{u}^{d=1} - \beta_0 - \beta_1 \bar{x}^{d=0} - \bar{u}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}} \Leftrightarrow$$

$$\tilde{\beta}_1 = \frac{\beta_1 \bar{x}^{d=1} + \bar{u}^{d=1} - \beta_1 \bar{x}^{d=0} - \bar{u}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}} \Leftrightarrow$$

$$\tilde{\beta_1} = \frac{\beta_1 \bar{x}^{d=1} - \beta_1 \bar{x}^{d=1}}{\bar{x}^{d=1} - \bar{x}^{d=0}} + \frac{\bar{u}^{d=1} - \bar{u}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}} \Leftrightarrow$$

$$\tilde{\beta_1} = \frac{\beta_1(\bar{x}^{d=1} - \bar{x}^{d=1})}{\bar{x}^{d=1} - \bar{x}^{d=0}} + \frac{\bar{u}^{d=1} - \bar{u}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}} \Leftrightarrow$$

$$\tilde{\beta_1} = \beta_1 + \frac{\bar{u}^{d=1} - \bar{u}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}}$$

Vi har nu et udtryk, hvor vi kan teste for konsistens. Tager $p\lim$ på begge sider:

$$p\lim(\tilde{\beta_1}) = \beta_1 + p\lim(\frac{\bar{u}^{d=1} - \bar{u}^{d=0}}{\bar{x}^{d=1} - \bar{x}^{d=0}}) \Leftrightarrow$$

$$p\lim(\tilde{\beta_1}) = \beta_1 + p\lim(\frac{1}{\bar{x}^{d=1} - \bar{x}^{d=0}} \cdot \bar{u}^{d=1} - \bar{u}^{d=0}) \Leftrightarrow$$

$$p\lim(\tilde{\beta_1}) = \beta_1 + p\lim\left(\frac{1}{\bar{x}^{d=1} - \bar{x}^{d=0}} \cdot \left(\frac{\sum_{i=1}^n u_i d_i}{\sum_{i=1}^n d_i}\right) - \left(\frac{\sum_{i=1}^n u_i(1-d_i)}{\sum_{i=1}^n (1-d_i)}\right)\right) \Leftrightarrow$$

$$p\lim(\tilde{\beta_1}) = \beta_1 + p\lim\left(\frac{1}{\bar{x}^{d=1} - \bar{x}^{d=0}} \cdot \left(\frac{\frac{1}{n}\sum_{i=1}^n u_i d_i}{\frac{1}{n}\sum_{i=1}^n d_i}\right) - \left(\frac{\frac{1}{n}\sum_{i=1}^n u_i(1-d_i)}{\frac{1}{n}\sum_{i=1}^n (1-d_i)}\right)\right) \Leftrightarrow$$

$$p\lim(\tilde{\beta_1}) = \beta_1 + p\lim\left(\frac{1}{\bar{x}^{d=1} - \bar{x}^{d=0}} \cdot \left(\frac{E(ud)}{\frac{1}{n}\sum_{i=1}^n d_i}\right) - \left(\frac{E(u) - E(ud)}{\frac{1}{n}\sum_{i=1}^n (1-d_i)}\right)\right) \Leftrightarrow$$

$$p\lim(\tilde{\beta_1}) = \beta_1 + p\lim\left(\frac{1}{\bar{x}^{d=1} - \bar{x}^{d=0}} \cdot \left(\frac{0}{\frac{1}{n}\sum_{i=1}^n d_i}\right) - \left(\frac{0-0}{\frac{1}{n}\sum_{i=1}^n (1-d_i)}\right)\right) \Leftrightarrow$$

$$p\lim(\tilde{\beta_1}) = \beta_1$$

Det er hermed vist, at estimatoren for $\tilde{\beta_1}$ er konsistent estimator for β_1.

Opgave 5

(a) Vi opstiller et simulationseksperiment med 500 observationer kørt 1000 gange. Vi anvender seed 117. Den efterspurgte deskriptive statistik er anført i tabel 3 og histogrammet figur 2 (se sidste afdeling af opgaven). Ud fra resultaterne kan det udledes, at OLS-estimatet i gennemsnit er -0,323 og standardfejlen er 0,0128

(b) I simulationen er både de robuste og ikke robuste standardfejl beregnet for β_1. Disse er anført i tabel 3. Det kan udledes, at de ikke robust standardfejl er 0,0153. Dette er et stykke fra standardfejlen på OLS-estimatoren, som er beregnet til 0,0114. Kigger vi derimod på de robuste standardfejl af OLS, som er på 0,0131, er vi tættere på β_1s standardfejl. Fejlen ved ikke at bruge robuste standardfejl er, at man kan få for store standardfejl, hvilket vil give konsekvenser ift. teststørrelsen, som giver problemer ved sandsynligheden for at forkaste en sand nulhypotsese.

(c) Vi kigger på forskellen mellem følgende to regressioner:

$$\gamma_0 + \gamma_1 * 1 + u$$

$$\gamma_0 + \gamma_1 * 0 + u$$

Dette kan opskrives som:

$$P(y=1|x=1) - P(y=1|x=1) = 1 - \Phi(-\gamma_0 - \gamma_1) - (1 - \Phi(-\gamma_0)) \Leftrightarrow$$

Indsætter værdierne givet for $\gamma_0\gamma_1$

$$P(y=1|x=1) - P(y=1|x=1) = -\Phi(-3-(-2)) + \Phi(-3) \Leftrightarrow$$

$$P(y=1|x=1) - P(y=1|x=1) = \Phi(-3) - \Phi(1) \Leftrightarrow$$

$$P(y=1|x=1) - P(y=1|x=1) = -0,157$$

Dette sammenlignes med OLS-estimatoren -0,32. Den store forskel er pga. i 5.c antager vi en konstant størrelse af x, hvor i modellen lader vi x frit flyde.

Hjælpesætninger, tabeller og figurer

Der benyttes følgende hjælpsætninger for summer i opgaven:

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} y_i(x_i - \bar{x})$$

$$\sum_{i=1}^{n}(x_i - \bar{x}^2) = \sum_{i=1}^{n} x_i(x_i - \bar{x})$$

Nedenunder kan relevante figurer og tabeller observeres:

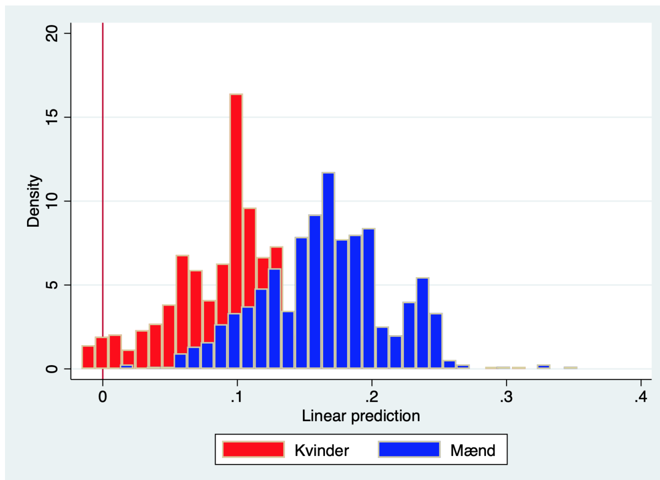

Figure 1: Histogram over prædikterede sandsynligheder

	mean	sd	min	max
OLS estimat	-0.3233	0.0128	-0.370	-0.275
SE OLS	0.0153	0.0006	0.013	0.018
Robust SE OLS	0.0131	0.0009	0.011	0.017
No. replikationer	1000			

Table 3: Resultater af simulationen

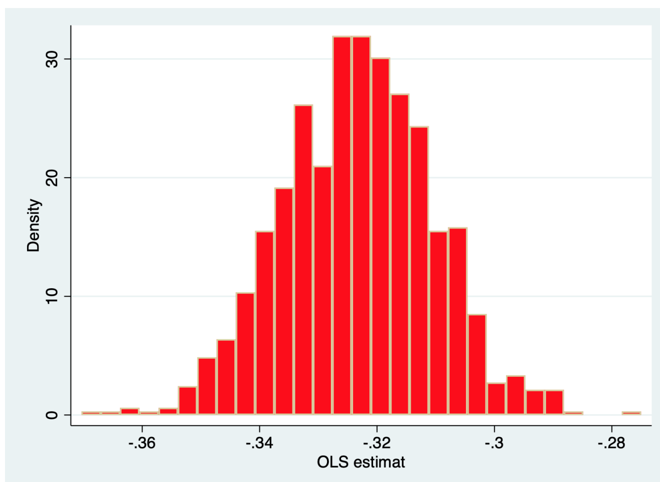

Figure 2: Fordeling over β_1 værdier.