

**NTNU, TRONDHEIM**

Norges teknisk-naturvitenskapelige universitet

Institutt for sosiologi og statsvitenskap

**EKSAMENSOPPGÅVE  
I  
SVSOS316  
REGRESJONSANALYSE**

Eksamensdag: 22 mai 2001

Eksamensstad: Dragvoll

Tid til eksamen: 6 timar

Vekter: 5

Talet av sider med tekst: 52

Talet av sider bokmål: -

Talet av sider nynorsk: 4

Talet av sider engelsk: -

Talet av sider vedlegg: 48

Dato for sensur:

Hjelpemiddel som kan nyttast:

Kalkulator.

Norsk-engelsk/ engelsk-norsk ordbok.

Hamilton, Lawrence C. 1992 «Regression with Graphics», Belmont, Duxbury,  
Hardy, Melissa A. 1993 «Regression with Dummy Variables», QASS 93, London,  
Sage

Breen, Richard 1996 «Regression Models. Censored, Sample Selected, or  
Truncated Data», QASS 111, London, Sage

---

**OPPGÅVE 1** (vekt 0,1)

- a) Forklar kva glatting (smoothing) er.
- b) Korleis kan ein nytte dummyvariablar til å teste for kurvesamanhengar?

**OPPGÅVE 2** (OLS-regresjon, vekt 0,45)

I tabellvedlegget til oppgåve 2 er det estimert 8 modellar av eiga inntekt (E.inntekt). Modellane 1-6 nyttar estimerte verdiar for manglande inntektsopplysningar, modellane 7 og 8 utelet personane der opplysningar mangla.

- a) Bruk modell 1 for å finne eit konfidensintervall for effekten av å ha heiltidsarbeid. Vurder om det er ein lineær eller kurvelineær samanheng mellom alder og inntekt. Finn ut frå modell 3 forventa inntekt for ei 40 år gammal kvinne med 12 års utdanning og heiltidsarbeid ved NTNU.
  - b) Formuler den modellen som er estimert som Modell 4. Vurder om testane i modell 4 er truverdige. Test om bustad gir ei signifikant yting til å forklare variasjonen i inntekt.
  - c) Modell 6 er identisk med modell 4 men er estimert utan den personen som har størst innverknad på estimatet av modell 4. Kva kan seiast om denne personen? Kva konsekvensar har det for regresjonsresultatet at personen vert utelaten?
  - d) Manglande opplysningar i variabelen "E.inntekt m/ est. missing" er erstatta med eit estimat frå Modell 7. Modell 7 er identisk med modell 3 men er estimert på faktiske observasjonar. Drøft generelt problemet med manglande observasjonar på avhengig variabel. Vurder konkret og substansielt skilnadene mellom Modell 3 og 7 i estimerte effektar for "Mann" og "Offentleg sektor".
-

**OPPGÅVE 3** (Logistisk regresjon, vekt 0,45)

I tabellvedlegget til oppgave 3 er det estimert 4 ulike modeller av "Besøke lokalt kunstgalleri"

- a) Lag eit konfidensintervall for effekten av "Mors utdanning" i modell 1. Korleis tolkar ein parameterestimaten for "Mors utdanning"?
  - b) Formuler den modellen som er estimert i modell 3. Finn ut om "Bustadstype" gir eit signifikant bidrag til modellen. Vurder om føresetnadene for modellen kan seiast å vere stetta.
  - c) Bruk modell 4 til å finne forventa verdi av sannsynet for å vitje det lokale kunstgalleriet for ein 50 år gammal mannleg universitetslærer frå Trondheim med 19 års utdanning når du også får vite at mor hans hadde 8 års utdanning. Skriv opp formelen for å finne betinga effektplott for samanhengen mellom sannsyn og alder.
  - d) Kva er definisjonen av Odds for å vitje lokalt kunstgalleri for den persontypen som er definert i pkt c)? Bruk definisjonen og modell 3 til å finne oddsraten for å velje å vitje lokalt kunstgalleri mellom ein mann med 19 års utdanning og ein med 18 års utdanning.
-



**OPPGAVE 2 Variabeldefinisjoner/ OPPGAVE 2 Variabeldefinisjoner/****QUESTION 2 Definition of variables**

**E.inntekt** = bruttoinntekt til intervjupersonen i 1000 kr, er koda

30, dersom personen seier bruttoinntekta er i intervallet 0-60.000

80, dersom personen seier bruttoinntekta er i intervallet 60-100.000

115, dersom personen seier bruttoinntekta er i intervallet 100-130.000

140, dersom personen seier bruttoinntekta er i intervallet 130-160.000

180, dersom personen seier bruttoinntekta er i intervallet 160-200.000

250, dersom personen seier bruttoinntekta er i intervallet 200-300.000

350, dersom personen seier bruttoinntekta er i intervallet 300-400.000

450, dersom personen seier bruttoinntekta er i intervallet 400.000+...

998, dersom personen ikkje vil svare

999, dersom opplysningar manglar

kodane 998 og 999 er i somme modellar erstatta med eit regresjonsestimat basert på

dei personane som har valide inntektsdata

**E.inntekt** = bruttoinntekt til intervjupersonen i 1000 kr, er kodet

30, hvis personen sier bruttoinntekta er i intervallet 0-60.000

80, hvis personen sier bruttoinntekta er i intervallet 60-100.000

115, hvis personen sier bruttoinntekta er i intervallet 100-130.000

140, hvis personen sier bruttoinntekta er i intervallet 130-160.000

180, hvis personen sier bruttoinntekta er i intervallet 160-200.000

250, hvis personen sier bruttoinntekta er i intervallet 200-300.000

350, hvis personen sier bruttoinntekta er i intervallet 300-400.000

450, hvis personen sier bruttoinntekta er i intervallet 400.000+...

998, hvis personen ikke vil svare

999, hvis opplysninger mangler

kodene 998 og 999 er i noen modeller erstattet med et regresjonsestimat basert på

de personene som har valide inntektsdata

**E.inntekt** (own income) = gross income of interviewee in 1000 kr, has been coded

60, if the person says the gross income is in the range of 0-60.000

90, if the person says the gross income is in the range of 60-100.000

120, if the person says the gross income is in the range of 100-130.000

150, if the person says the gross income is in the range of 130-160.000

180, if the person says the gross income is in the range of 160-200.000

250, if the person says the gross income is in the range of 200-300.000

350, if the person says the gross income is in the range of 300-400.000

450, if the person says the gross income is in the range of 400.000+...

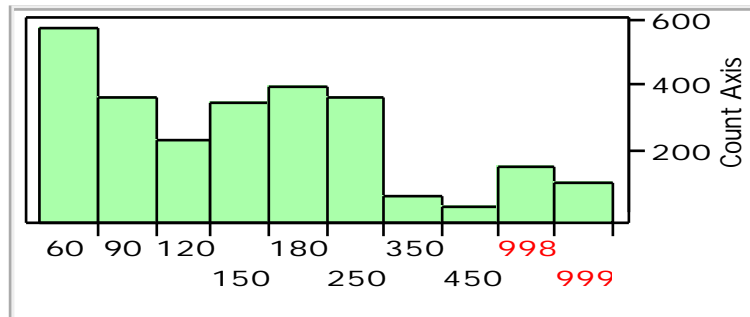
998, if the person will not answer

999, if data is missing

the codes 998 and 999 are in some models replaced by a regression estimate based on

the persons with valid income data

**OPPGÅVE 2 Variabeldefinisjoner/ OPPGÅVE 2 Variabeldefinisjoner/  
QUESTION 2 Definition of variables**



**Frequencies**

Income level	Count	Prob
60	613	0.20794
90	396	0.13433
120	261	0.08853
150	379	0.12856
180	435	0.14756
250	399	0.13535
350	94	0.03189
450	57	0.01934
998= unwilling to answer	179	0.06072
999= missing	135	0.04579
Total	2948	1.00000

**E.inntekt**

**missing utelaten/ utelatt/ excluded**

**Frequencies**

Income Level	Count	Prob
60	613	0.23273
90	396	0.15034
120	261	0.09909
150	379	0.14389
180	435	0.16515
250	399	0.15148
350	94	0.03569
450	57	0.02164
Total	2634	1.00000

**Mann**

Man

**Frequencies**

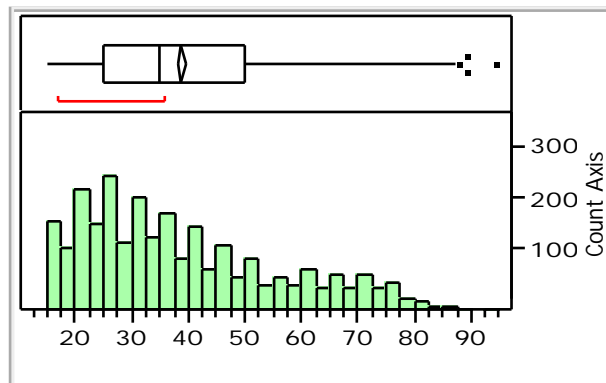
Level	Count	Prob
0	1503	0.50984
1 = mann (man)	1445	0.49016
Total	2948	1.00000

**OPPGÅVE 2 Variabeldefinisjoner/ OPPGAVE 2 Variabeldefinisjoner/  
QUESTION 2 Definition of variables****Heiltidsarbeid** = 1, ja,arbeider heil dag; 0 = andre svar (deltid, varierer, nei, missing)**Heiltidsarbeid** = 1, ja, arbeider hel dag; 0 = andre svar (deltid, varierer, nei, missing)**Heiltidsarbeid** = 1, yes, full time employment; 0=other answers (part time, varying, no, missing)**Frequencies**

Level	Count	Prob
0	1669	0.56615
1 = heiltid (full time)	1279	0.43385
Total	2948	1.00000

**Offentleg sektor** = 1, arbeider i offentlig sektor; 0 = andre svar (privat, andre svar, missing)**Offentleg sektor** = 1, arbeider i offentlig sektor; 0 = andre svar (privat, andre svar, missing)**Offentleg sektor** = 1, employed by public sector; 0 = other answers (privat, other answers, missing)**Frequencies**

Level	Count	Prob
0	2131	0.72286
1 = ansatt (employed)	817	0.27714
Total	2948	1.00000

**Alder** = alder i år**Alder** = age in years**Quantiles**

100.0%	maximum	94.000
75.0%	quartile	50.000
50.0%	median	35.000
25.0%	quartile	25.000
0.0%	minimum	15.000

**Moments**

Mean	38.88501
Std Dev	17.44791
Std Err Mean	0.32135
upper 95% Mean	39.51511
lower 95% Mean	38.25490
N	2948

**OPPGÅVE 2 Variabeldefinisjoner/ OPPGAVE 2 Variabeldefinisjoner/****QUESTION 2 Definition of variables**

**E. utdanning** = «eiga utdanning i år» er koda

7 dersom personen har utdanning på folkeskolenivå (inntil 8 års skolegang)

9 dersom personen har utdanning på ungdomsskole/ realskolenivå (9-10 års skolegang)

12 dersom personen har utdanning på videregående skole/ gymnas nivå (11-13 års skolegang)

14 dersom personen har utdanning på universitetsnivå (meir enn 12 års skolegang + studiar), men er uten akademisk tittel

17 dersom personen har utdanning på universitetsnivå (meir enn 12 års skolegang + studiar), og har akademisk tittel

**E. utdanning** = «egen utdanning i år» er kodet

7 hvis personen har utdanning på folkeskolenivå (inntil 8 års skolegang)

9 hvis personen har utdanning på ungdomsskole/ realskolenivå (9-10 års skolegang)

12 hvis personen har utdanning på videregående skole/ gymnas nivå (11-13 års skolegang)

14 hvis personen har utdanning på universitetsnivå (mer enn 12 års skolegang + studier), men er uten akademisk tittel

17 hvis personen har utdanning på universitetsnivå (mer enn 12 års skolegang + studier), og har akademisk tittel

**E. utdanning** = «own education in years» has been coded

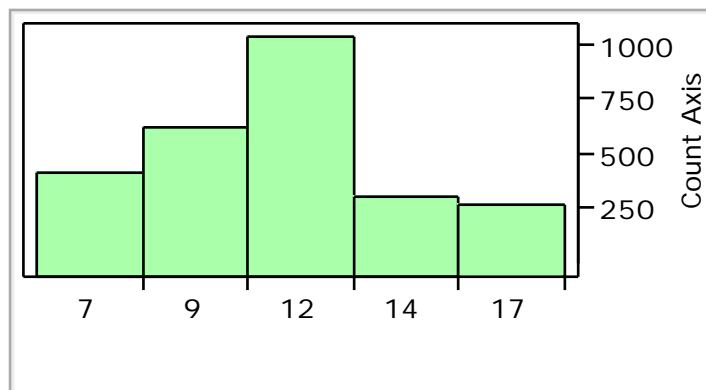
7 if the person has education at the level of primary school (up to 8 years of education)

9 if the person has education at the level of secondary school (9-10 years of education)

12 if the person has education at the level of high school (11-13 years of education)

14 if the person has education at college level (more than 12 years of education + college studies), but without academic title

17 if the person has education at college level (more than 12 years of education + college studies), and with academic title

**Frequencies**

Years	Count	Prob
7	475	0.16113
9	680	0.23066
12	1092	0.37042
14	371	0.12585
17	330	0.11194
Total	2948	1.00000



**OPPGÅVE 2 Variabeldefinisjoner/ OPPGAVE 2 Variabeldefinisjoner/  
QUESTION 2 Definition of variables****Bustadstype** = bustadstype

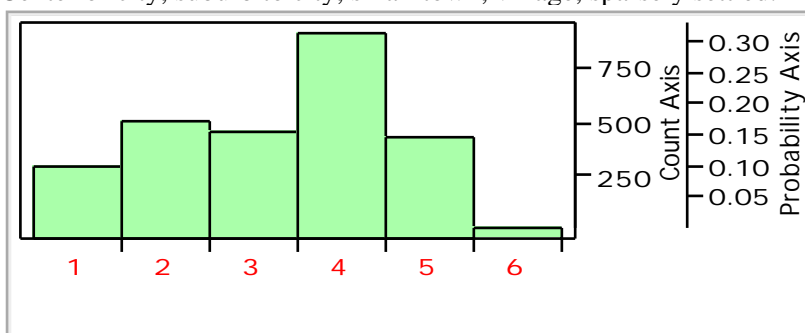
er svar på spørsmålet: Kva for ein av dei omtala bustadane liknar mest på den du har i dag? Sentrum av storby, Forstad til storby, Småby, Tettstad i landkommune eller Spredtbygd.

**Bustadstype** = bostedstype

er svar på spørsmålet: «Hvilke av de beskrevne bostedene ligner mest på det du har i dag? Sentrum av storby, Forstad til storby, Småby, Tettsted i landkommune eller Spredtbygd»

**Bustadstype** = type of home town

is the answer to the question Which of the described places resembles closest where you live today? Center of city, suburb to city, small town, village, sparsely settled.

**Frequencies**

Type	Count	Prob
1 = sentrum av storby (center of city)	354	0.12008
2 = forstad til storby (suburb of city)	562	0.19064
3 = småby (small town)	510	0.17300
4 = tettstad i landkommune (village)	969	0.32870
5 = spredtbygd (sparsely settled)	492	0.16689
6 = ukjent (unknown)	61	0.02069
Total	2948	1.00000

**Bost. type** er dummykoda med**Bost. Sentrum storby**

=1 dersom bustaden ligg i sentrum av storby, 0 elles

**Bost. Forst storby**

=1 dersom bustaden ligg i forstad til storby, 0 elles

**Bost. Småby**

=1 dersom bustaden ligg i småby, 0 elles

**Bost. Tettst**

=1 dersom bustaden ligg i tettstad i landkommune, 0 elles

**Bost. Spredtb**

=1 dersom bustaden ligg i spredtbygde strok, 0 elles

**Bost. Uoppg**

=1 dersom det ikkje er oppgitt kvar bustaden ligg, 0 elles

**Bost. type** er dummykodet med**Bost. Sentrum storby**

=1 hvis bostedet ligger i sentrum av storby, 0 ellers

**Bost. Forst storby**

=1 hvis bostedet ligger i forsted til storby, 0 ellers

**Bost. Småby**

=1 hvis bostedet ligger i småby, 0 ellers

**Bost. Tettst**

=1 hvis bostedet ligger i tettsted i landkommune, 0 ellers

**Bost. Spredtb**

=1 hvis bostedet ligger i spredtbygd område, 0 ellers

**Bost. Uoppg**

=1 hvis det ikke er oppgitt hvor bostadet ligger, 0 ellers

**Bost.type** has been dummy coded with**Bost. Sentrum storby**

=1 if the home town is center of city, 0 otherwise

**Bost. Forst storby**

=1 if the home town is suburb of city, 0 otherwise

**Bost. Småby**

=1 if the home town is small town, 0 otherwise

**Bost. Tettst**

=1 if the home town is village, 0 otherwise

**Bost. Spredt**

=1 if the home town is sparsely settled, 0 otherwise

**Bost. Uoppg**

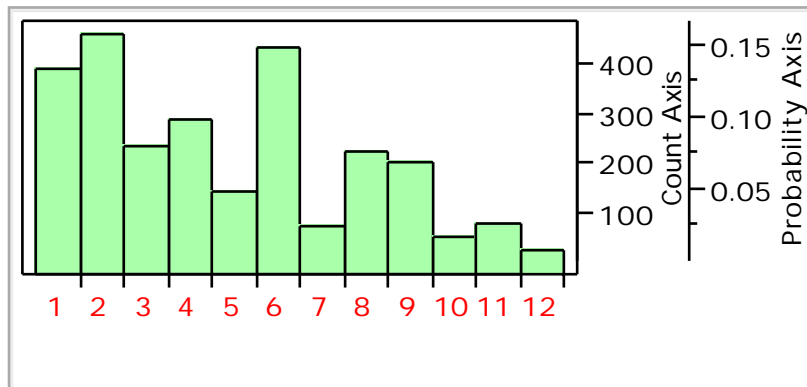
=1 if the home town is unknown, 0 otherwise

## OPPGÅVE 2 Variabeldefinisjoner/ OPPGAVE 2 Variabeldefinisjoner/ QUESTION 2 Definition of variables

**Kjelde til livsopphald** = kjelde til livsopphald

**Kjelde til livsopphald** = kilde til livsopphold

**Kjelde til livsopphald** = source of livelihood



### Frequencies

Code	Count	Prob
1 Arbeidar, ufaglært (Worker, unskilled)	420	0.14247
2 Arbeidar, faglært (Worker, skilled)	489	0.16588
3 Funksjonær, leiande, (White collar worker, manager)	264	0.08955
4 Funksjonær elles (White collar worker otherwise)	320	0.10855
5 Sjølvstendig (Self employed)	171	0.05801
6 Elev/ student/ lærling (Student)	463	0.15706
7 For tida arbeidslaus (Currently unemployed)	99	0.03358
8 Alderspensjonist (Pensioner)	255	0.08650
9 Trygda (Social security)	232	0.07870
10 Gift utan betalt arbeid (Married without paid work)	79	0.02680
11 Andre kjelder (Other sources)	107	0.03630
12 Uoppgitt (No answer)	49	0.01662
Total	2948	1.00000

**Kjelde til livsopphald** er dummykoda med

<b>Arbeider</b>	=1 dersom personen er faglært eller ufaglært arbeidar (kode 1 og 2), 0 elles
<b>Funksjonær</b>	=1 dersom personen er funksjonær (inkl leiande funksjonær) (kode 3 og 4), 0 elles
<b>Selvstendig</b>	=1 dersom personen er sjølvstendig (kode 5), 0 elles
<b>Elev/ student</b>	=1 dersom personen er elev, student eller lærling (kode 6), 0 elles
<b>Pensjon/ trygd</b>	=1 dersom personen er alderspensjonist eller trygda (kode 8 og 9), 0 elles
<b>Annet</b>	=1 dersom personen er arbeidslaus, gift utan eige betalt arbeid eller har anna type forsørging (kode 7, 10 og 11), 0 elles
<b>Uoppgitt KtL</b>	=1 dersom personen ikkje har svart på spørsmålet, 0 elles

**OPPGÅVE 2 Variabeldefinisjoner/ OPPGAVE 2 Variabeldefinisjoner/  
QUESTION 2 Definition of variables**

**Kjelde til livsopphald** er dummykodet med

<b>Arbeider</b>	=1 hvis personen er faglært eller ufaglært arbeider (kode 1 og 2), 0 ellers
<b>Funksjonær</b>	=1 hvis personen er funksjonær (inkl ledende funksjonær) (kode 3 og 4), 0 ellers
<b>Selvstendig</b>	=1 hvis personen er selvstendig (kode 5), 0 ellers
<b>Elev/ student</b>	=1 hvis personen er elev, student eller lærling (kode 6), 0 ellers
<b>Pensjon/ trygd</b>	=1 hvis personen er alderspensjonist eller trygdet (kode 8 og 9), 0 ellers
<b>Annet</b>	=1 hvis personen er arbeidsløs, gift uten ege tbetalt arbeid eller har annen type forsørging (kode 7, 10 og 11), 0 ellers
<b>Uoppgitt KtL</b>	=1 hvis personen ikke har svart på spørsmålet, 0 ellers

**Kjelde til livsopphald** has been dummy coded with

<b>Arbeider</b>	=1 if the person is skilled or unskilled worker (code 1 og 2), 0 otherwise
<b>Funksjonær</b>	=1 if the person is white collar worker or manager (code 3 og 4), 0 otherwise
<b>Selvstendig</b>	=1 if the person is selfemployed (code 5), 0 otherwise
<b>Elev/ student</b>	=1 if the person is student (code 6), 0 otherwise
<b>Pensjon/ trygd</b>	=1 if the person is pensioner (code 8 og 9), 0 otherwise
<b>Annet</b>	=1 if the person is unemployed, married without paid work or has other source of livelihood (code 7, 10 og 11), 0 otherwise
<b>Uoppgitt KtL</b>	=1 if the person did not answer the question, 0 otherwise

---

**OPPGAVE 2 / OPPGAVE 2 Modell 1 /QUESTION 2 Model 1**

Response E.innt m/ est. missing

**Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.513328
RSquare Adj	0.512501
Root Mean Square Error	61.91234
Mean of Response	141.5834
Observations (or Sum Wgts)	2948

**Analysis of Variance**

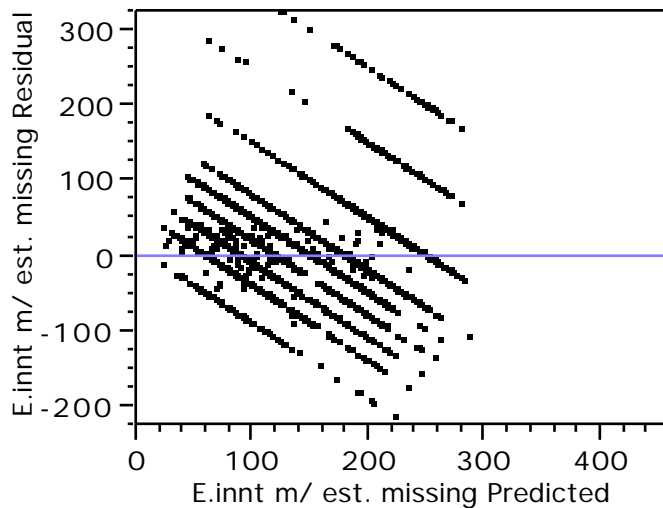
The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	5	11894766	2378953	620.6282	
Error	2942	11277090	3833		<b>Prob &gt; F</b>
C. Total	2947	23171856			0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	-46.72458	6.088175	-7.67	<.0001	.
Alder	1.1407587	0.068913	16.55	<.0001	1.1115088
Mann	31.209589	2.3943	13.03	<.0001	1.1017984
E.utdanning	7.5193284	0.416488	18.05	<.0001	1.1920058
Heiltidsarbeid	98.818511	2.499229	39.54	<.0001	1.1799326
Offentleg sektor	2.551218	2.705888	0.94	0.3458	1.1280887

**Residual by Predicted Plot**



**OPPGÅVE 2/ OPPGÅVE 2 Modell 2 /QUESTION 2 Model 2**

Response E.innt m/ est. missing

**Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.554366
RSquare Adj	0.553457
Root Mean Square Error	59.25457
Mean of Response	141.5834
Observations (or Sum Wgts)	2948

**Analysis of Variance**

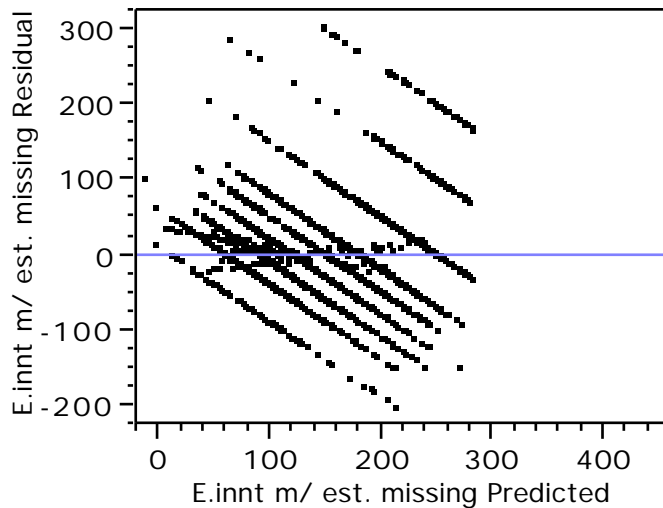
The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	6	12845699	2140950	609.7655	
Error	2941	10326157	3511		
C. Total	2947	23171856			0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	-144.4565	8.319775	-17.36	<.0001	.
Alder	6.9577419	0.359564	19.35	<.0001	33.035148
Mann	35.060961	2.303437	15.22	<.0001	1.11329
E.utdanning	7.2228987	0.399016	18.10	<.0001	1.1944399
Heiltidsarbeid	83.17854	2.57382	32.32	<.0001	1.3661933
Offentleg sektor	-6.262291	2.644524	-2.37	0.0179	1.1763307
Alder*Alder	-0.064837	0.00394	-16.46	<.0001	33.921502

**Residual by Predicted Plot**



**OPPGÅVE 2/ OPPGAVE 2 Modell 3 /QUESTION 2 Model 3**

Response E.innt m/ est. missing

**Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.576719
RSquare Adj	0.575133
Root Mean Square Error	57.79855
Mean of Response	141.5834
Observations (or Sum Wgts)	2948

**Analysis of Variance**

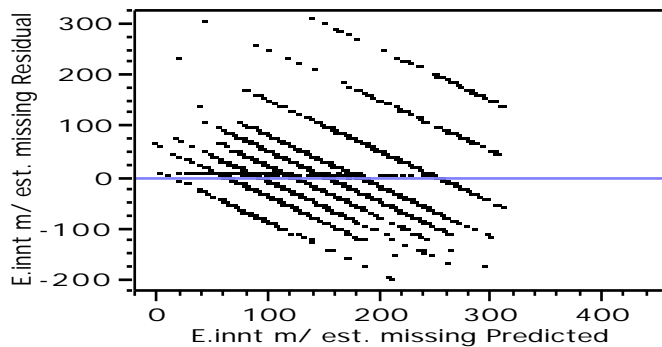
The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	11	13363643	1214877	363.6624	
Error	2936	9808213	3341		<b>Prob &gt; F</b>
C. Total	2947	23171856			0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	-68.01107	11.53876	-5.89	<.0001
Alder	4.3638167	0.476325	9.16	<.0001
Mann	-105.888	16.16992	-6.55	<.0001
E.utdanning	5.5702413	0.588171	9.47	<.0001
Heiltidsarbeid	66.897826	3.468928	19.28	<.0001
Offentleg sektor	9.1808397	3.47846	2.64	0.0084
Alder*Alder	-0.041206	0.005215	-7.90	<.0001
Alder*Mann	4.5283472	0.71791	6.31	<.0001
Alder*Alder*Mann	-0.040529	0.007874	-5.15	<.0001
E.utdanning*Mann	3.0419769	0.786175	3.87	0.0001
Heiltidsarbeid*Mann	29.348031	5.146062	5.70	<.0001
Offentleg sektor*Mann	-30.07997	5.20596	-5.78	<.0001

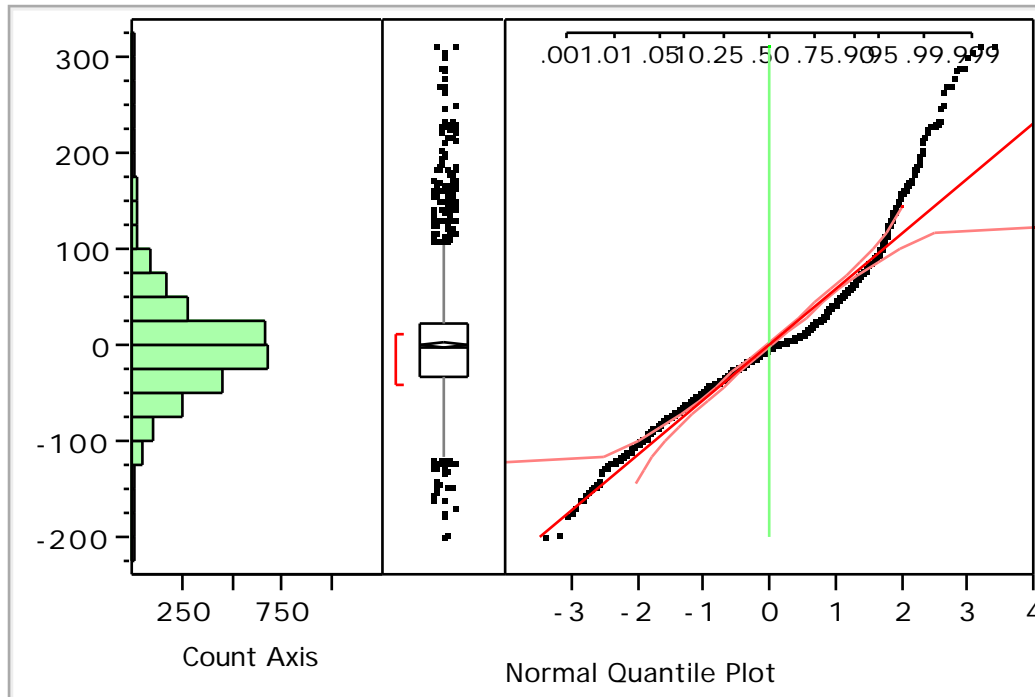
**Residual by Predicted Plot**



**OPPGAVE 2/ OPPGAVE 2 Modell 3 /QUESTION 2 Model 3**

**Distributions**

**Residual E.innt(EM) Modell 3**



**Quantiles**

100.0%	maximum	312.16
99.5%		237.34
97.5%		151.18
90.0%		65.44
75.0%	quartile	22.37
50.0%	median	-3.49
25.0%	quartile	-33.50
10.0%		-61.59
2.5%		-100.37
0.5%		-142.47
0.0%	minimum	-200.59

**Moments**

Mean	4.7e-14
Std Dev	57.690575
Std Err Mean	1.0625298
upper 95% Mean	2.0834131
lower 95% Mean	-2.083413
N	2948

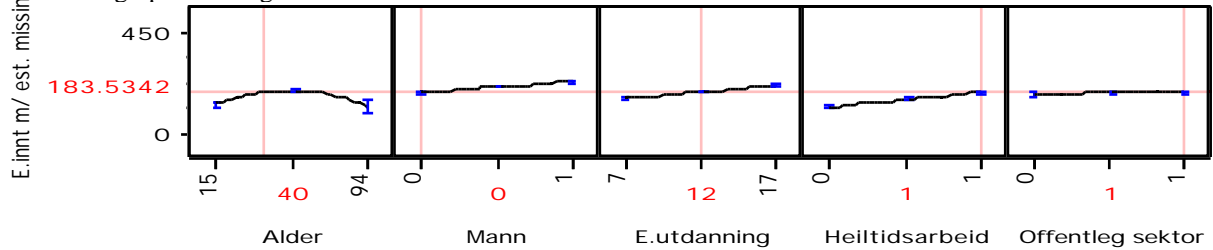
**OPPGAVE 2/ OPPGAVE 2 Modell 3 / QUESTION 2 Model 3**

Response E.innt m/ est. missing

**Prediction Profile<sup>1</sup>**

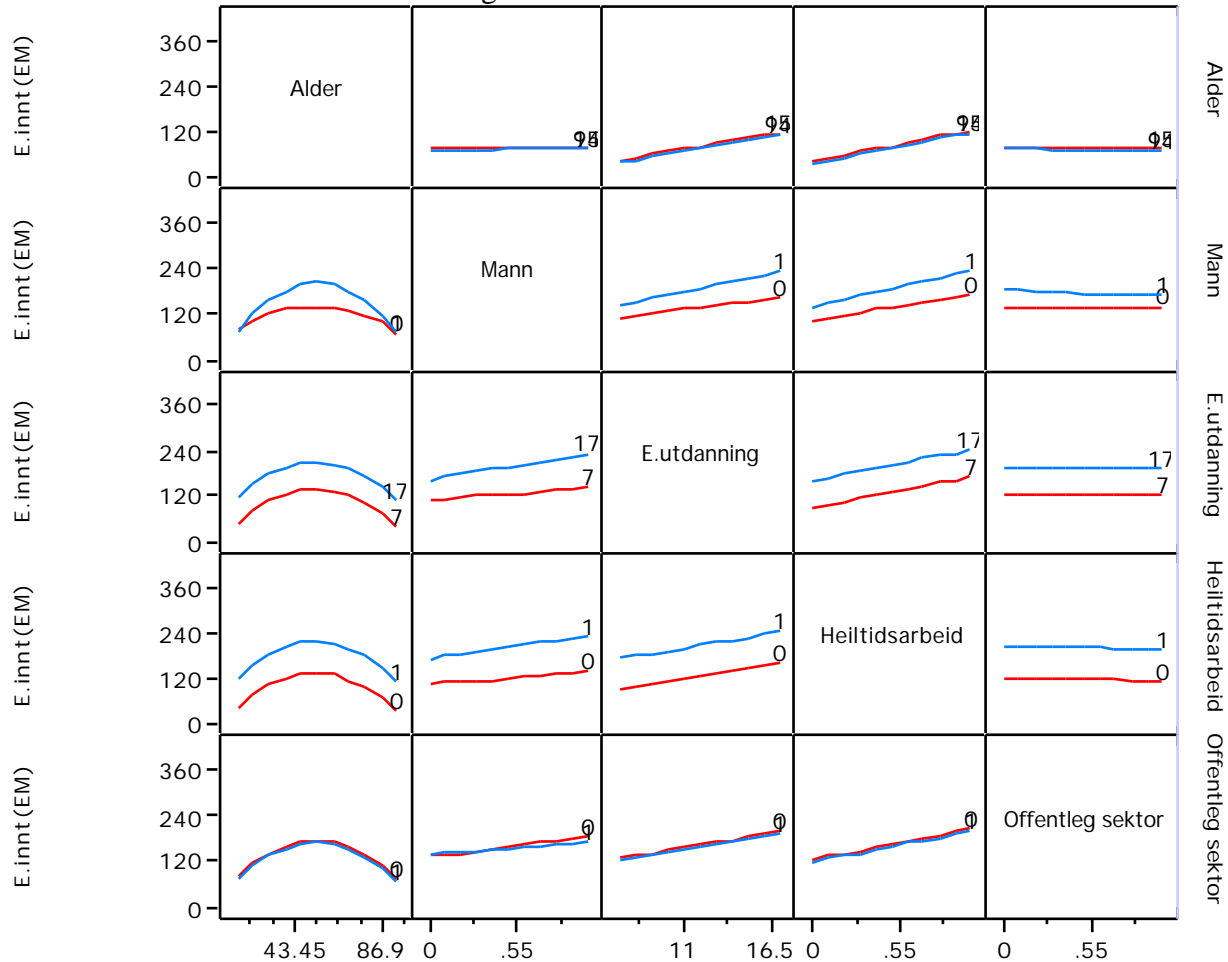
The change in predicted response as you vary one factor at a time, holding the other factors at their current values.

Click in the graph to change the current values of the factors.



**Interaction Profiles**

The change in predicted response by the variable named in the diagonal and for low and high values of the variable named on the right side of the row.



<sup>1</sup> Mer om lesing av plottet er gitt under vedlegg til oppgave 3. Mer om lesing av plottet er gitt under vedlegg til oppgave 3. More on the reading of the plot is presented in attachments to question 3



**OPPGÅVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

Response E.innt m/ est. missing

**Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.609627
RSquare Adj	0.606691
Root Mean Square Error	55.61059
Mean of Response	141.5834
Observations (or Sum Wgts)	2948

**Analysis of Variance**

The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

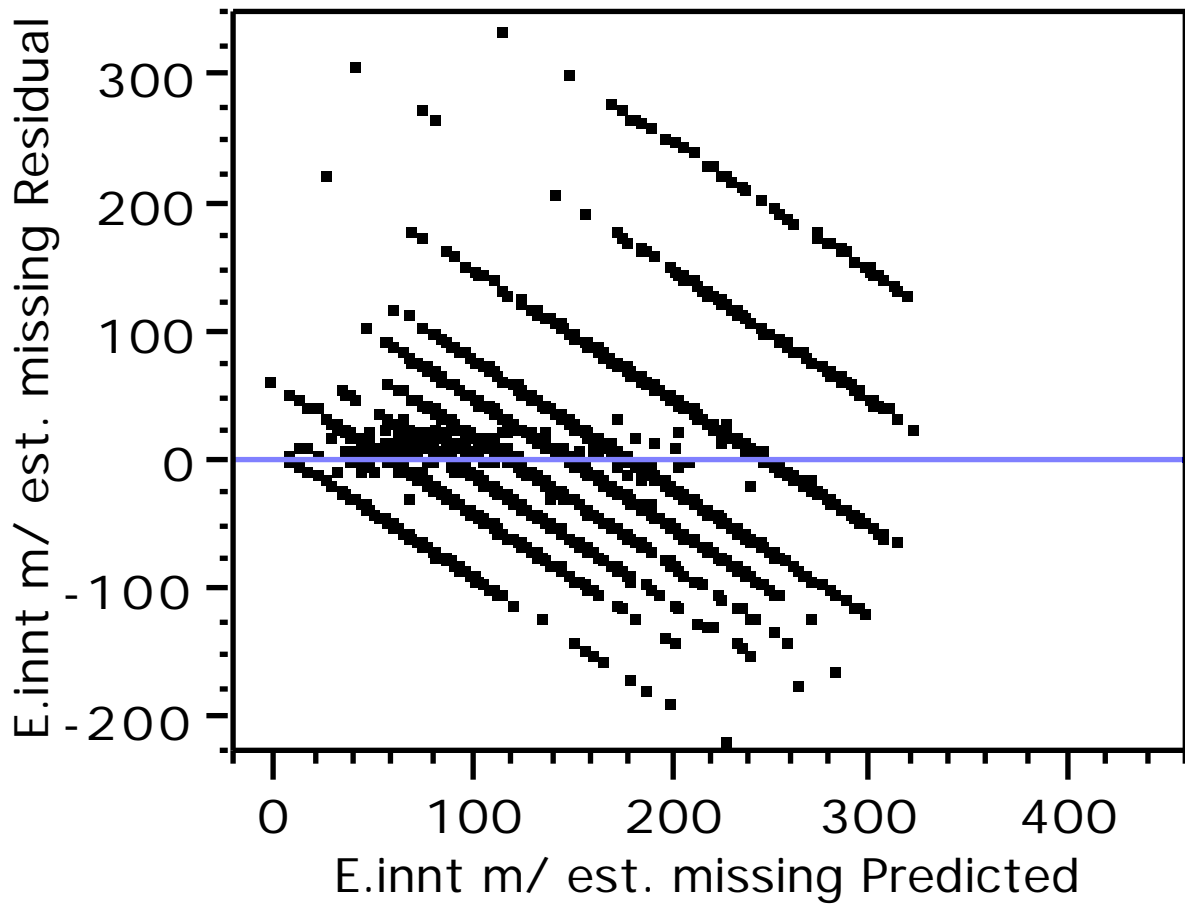
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	22	14126182	642099	207.6285	
Error	2925	9045674	3093		
C. Total	2947	23171856			0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	-25.75671	12.27589	-2.10	0.0360	.
Alder	2.7856639	0.493533	5.64	<.0001	70.66186
Mann	-98.63631	15.61582	-6.32	<.0001	58.091644
E.utdanning	4.4888481	0.582243	7.71	<.0001	2.8875032
Heiltidsarbeid	47.809716	3.645847	13.11	<.0001	3.1122989
Offentleg sektor	7.0106699	3.488949	2.01	0.0446	2.3246212
Alder*Alder	-0.026519	0.005383	-4.93	<.0001	71.909952
Alder*Mann	4.6389816	0.693943	6.68	<.0001	249.95149
Alder*Alder*Mann	-0.042205	0.007614	-5.54	<.0001	122.58306
E.utdanning*Mann	2.5075154	0.759597	3.30	0.0010	20.607265
Heiltidsarbeid*Mann	29.007274	5.046504	5.75	<.0001	4.8152575
Offentleg sektor*Mann	-26.08115	5.078068	-5.14	<.0001	2.4242041
Sentrum storby	15.885219	4.060826	3.91	<.0001	1.6609613
Forstad storby	20.133127	3.575067	5.63	<.0001	1.8798906
Småby	8.8718665	3.562002	2.49	0.0128	1.7304121
Tettstad	8.8304113	3.111744	2.84	0.0046	2.0367357
Uoppg bostad	4.8355522	7.583838	0.64	0.5238	1.1109953
Funksjonær	33.852976	3.140055	10.78	<.0001	1.4931118
Sjølvstendig	24.907939	4.858604	5.13	<.0001	1.2295676
Elev/ student	-23.62938	4.221299	-5.60	<.0001	2.2488325
Pensjon/ trygd	-5.323715	4.640714	-1.15	0.2514	2.8311808
Andre KtL	-15.57144	4.11048	-3.79	0.0002	1.406561
Uoppgitt KtL	-17.46346	8.461389	-2.06	0.0391	1.1155398

**OPPGAVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

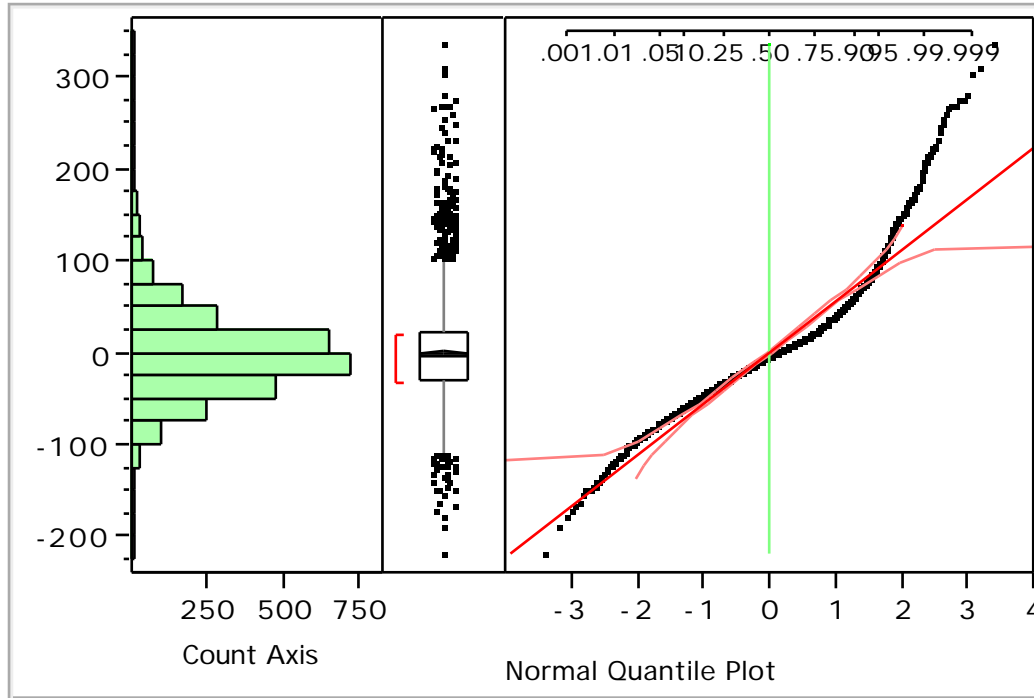
**Residual by Predicted Plot**



**OPPGAVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

**Distributions**

**Residual E.innt(EM) Modell 4**



**Quantiles**

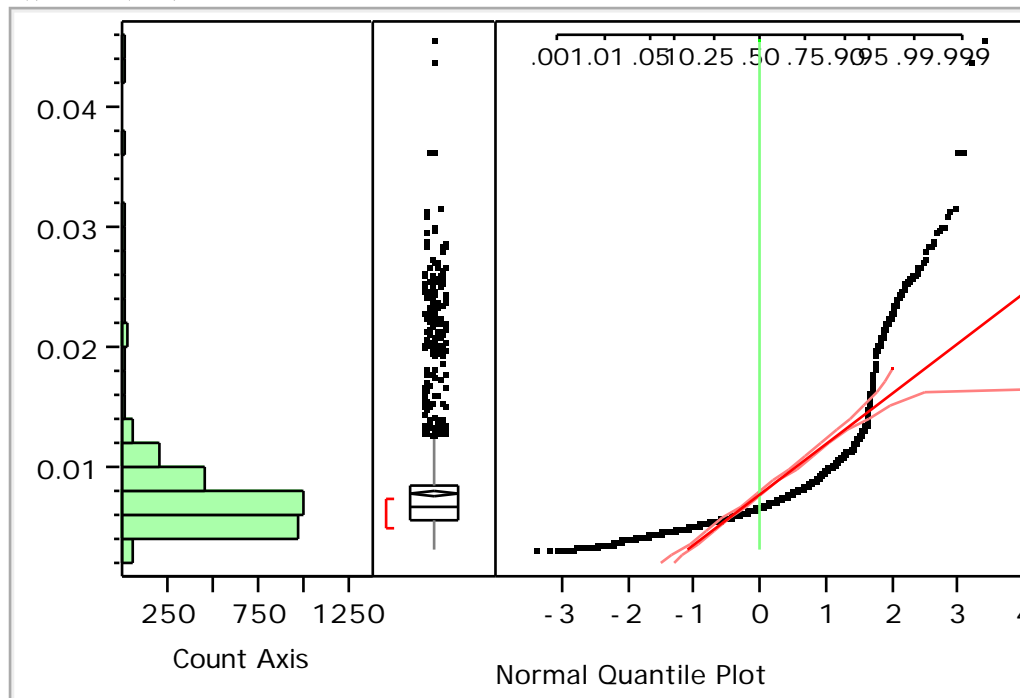
100.0%	maximum	336.37
99.5%		241.61
97.5%		145.99
90.0%		61.35
75.0%	quartile	21.89
50.0%	median	-4.63
25.0%	quartile	-31.06
10.0%		-59.15
2.5%		-92.04
0.5%		-140.36
0.0%	minimum	-218.00

**Moments**

Mean	8.9e-14
Std Dev	55.402632
Std Err Mean	1.0203911
upper 95% Mean	2.0007873
lower 95% Mean	-2.000787
N	2948

**OPPGAVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

**h(i) E.innt(EM) Modell 4**



**Quantiles**

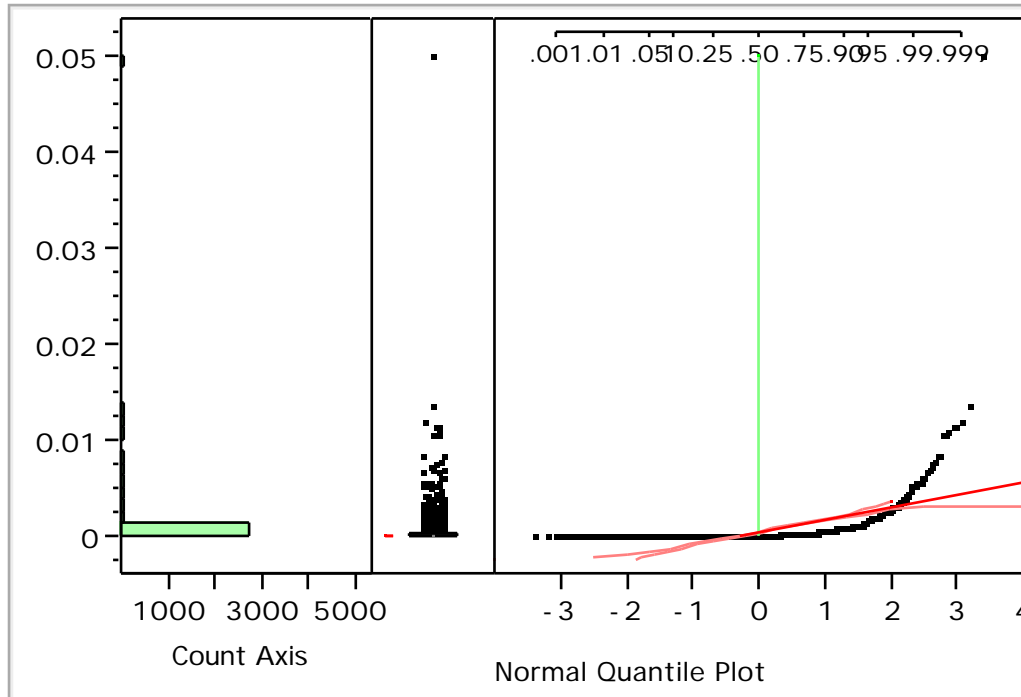
100.0%	maximum	0.04556
99.5%		0.02850
97.5%		0.02246
90.0%		0.01113
75.0%	quartile	0.00844
50.0%	median	0.00671
25.0%	quartile	0.00558
10.0%		0.00481
2.5%		0.00408
0.5%		0.00329
0.0%	minimum	0.00319

**Moments**

Mean	0.0078019
Std Dev	0.0041673
Std Err Mean	0.0000768
upper 95% Mean	0.0079524
lower 95% Mean	0.0076514
N	2948

**OPPGAVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

**Cook's D(i) E.innt(EM) Modell 4**



**Quantiles**

100.0%	maximum	0.04995
99.5%		0.00685
97.5%		0.00257
90.0%		0.00080
75.0%	quartile	0.00029
50.0%	median	0.00007
25.0%	quartile	0.00001
10.0%		0.00000
2.5%		0.00000
0.5%		0.00000
0.0%	minimum	0.00000

**Moments**

Mean	0.0003672
Std Dev	0.0013127
Std Err Mean	0.0000242
upper 95% Mean	0.0004147
lower 95% Mean	0.0003198
N	2948

**OPPGAVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

Case information for persons with high Cook's D. Sorted by Cood's D in Model 4

Case No i	Nær- ing	Ekte- skapeleg status	Heil- tids- arbeid	Offentleg sektor	Talet på arbeids- takarar i husst	Barn i hus- holdet	Mor's utdann ing	Far's utdann ing	E.utda ning
856	5	1	0	0	2	1	9	9	17
1733	5	1	1	1	2	1	12	15	17
1954	2	1	0	0	2	0	9	12	12
1756	6	1	0	1	2	0	7	7	17
2177	1	1	0	0	1	1	7	7	12
1329	2	1	0	0	2	1	7	7	14
69	6	2	1	1	1	0	12	7	17
1803	2	2	1	0	1	0	7	15	9
1706	2	1	1	0	2	1	7	7	12
1537	3	1	0	1	2	1	9	7	12

Næring=3=samferdsel/transport/post/tele; Ekteskapeleg status=1=gift; barn i husholdet=1=ja;

Case information for persons with high Cook's D. Sorted by Cood's D in Model 4

Case No i	Alder	Mann	E.inntekt (1000)	HH. inntekt (1000)	Busads-type	Kjelde til livsopphald
856	45	0	350	450	3	5
1733	34	1	450	450	3	11
1954	45	1	450	250	2	9
1756	57	0	450	450	2	4
2177	23	0	250	250	3	12
1329	40	1	450	450	2	5
69	22	1	450	450	2	5
1803	19	1	450	450	1	2
1706	39	1	450	450	6	5
1537	42	1	450	180	5	12

Case information for persons with high Cook's D. Sorted by Cood's D in Model 4

Case No i	Residual E.innt(EM) Modell 4	h(i) E.innt(EM) Modell 4	Cook's D(i) E.innt(EM) Modell 4	Residual E.innt(EM) Modell 3
856	194.013184	0.01526944	0.00833312	210.386585
1733	231.415303	0.01088822	0.0083793	194.295756
1954	280.684995	0.00937461	0.01058107	285.916767
1756	265.827868	0.01046762	0.01062053	299.275619
2177	180.440558	0.02287651	0.01096772	172.598158
1329	244.375995	0.01308455	0.01127906	278.416125
69	222.587646	0.01596968	0.01148784	246.076355
1803	302.464761	0.00919213	0.01204325	310.698154
1706	198.844729	0.02342504	0.01365383	201.829868
1537	336.365225	0.02957064	0.04994732	312.159773

**OPPGAVE 2/ OPPGAVE 2 Modell 4 /QUESTION 2 Model 4**

Case information for persons with high Cook's D. Sorted by Cood's D. Model 4

<b>Case No i</b>	<b>Residual E.innt Modell 7</b>	<b>Residual E.innt Modell 8</b>	<b>h(i) E.innt Modell 8</b>	<b>Cook's D(i) E.innt Modell 8</b>
856	202.339834	186.214926	0.01671544	0.00828783
1733	193.112349	219.940053	0.01243377	0.00852572
1954	282.643643	282.511693	0.01031839	0.01162375
1756	297.940583	265.159201	0.01124183	0.01117693
2177	165.767643	144.605977	0.04738446	0.01509474
1329	275.450862	242.111387	0.01392964	0.01160933
69	244.391694	221.266461	0.01678515	0.011752
1803	308.902315	302.600516	0.01007755	0.01301801
1706	201.671498	195.615925	0.02794339	0.01564432
1537	306.614555	304.13712	0.05274169	0.07516376

**OPPGAVE 2/ OPPGAVE 2 Modell 5 /QUESTION 2 Model 5**

Response E.innt m/ est. missing

**Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.604918
RSquare Adj	0.602626
Root Mean Square Error	55.8972
Mean of Response	141.5834
Observations (or Sum Wgts)	2948

**Analysis of Variance**

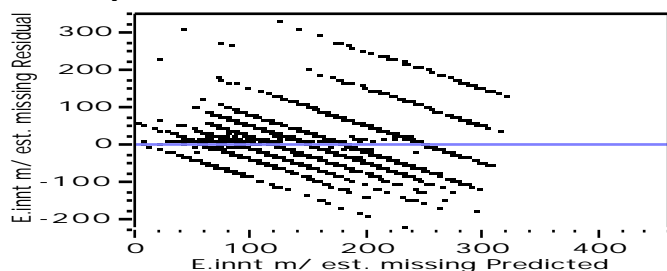
The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	17	14017080	824534	263.8934	
Error	2930	9154776	3124		<b>Prob &gt; F</b>
C. Total	2947	23171856			0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	-20.86695	12.17172	-1.71	0.0866	.
Alder	2.7388868	0.495241	5.53	<.0001	70.423981
Mann	-98.80523	15.68116	-6.30	<.0001	57.979621
E.utdanning	4.9697022	0.578128	8.60	<.0001	2.8177062
Heiltidsarbeid	49.584497	3.650571	13.58	<.0001	3.0884525
Offentleg sektor	6.1590024	3.502563	1.76	0.0788	2.3188352
Alder*Alder	-0.026003	0.0054	-4.82	<.0001	71.608844
Alder*Mann	4.6638045	0.697088	6.69	<.0001	249.64236
Alder*Alder*Mann	-0.042905	0.007648	-5.61	<.0001	122.43097
E.utdanning*Mann	2.5489381	0.762984	3.34	0.0008	20.578788
Heiltidsarbeid*Mann	27.615153	5.058318	5.46	<.0001	4.7883456
Offentleg sektor*Mann	-25.46626	5.101504	-4.99	<.0001	2.421607
Funksjonær	35.459941	3.140839	11.29	<.0001	1.4785777
Sjølvstendig	23.777823	4.868233	4.88	<.0001	1.2218195
Elev/ student	-22.19971	4.233252	-5.24	<.0001	2.2384531
Pensjon/ trygd	-4.88983	4.661647	-1.05	0.2943	2.8275595
Andre KtL	-14.93793	4.129228	-3.62	0.0003	1.4049024
Uoppgitt KtL	-16.55508	8.500319	-1.95	0.0516	1.1143131

**Residual by Predicted Plot**





**OPPGÅVE/ OPPGAVE 2 Modell 6 /QUESTION 2 Model 6**

Response E.innt m/ est. missing (case 1537 excluded)

**Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.613069
RSquare Adj	0.610158
Root Mean Square Error	55.2605
Mean of Response	141.4787
Observations (or Sum Wgts)	2947

**Analysis of Variance**

The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

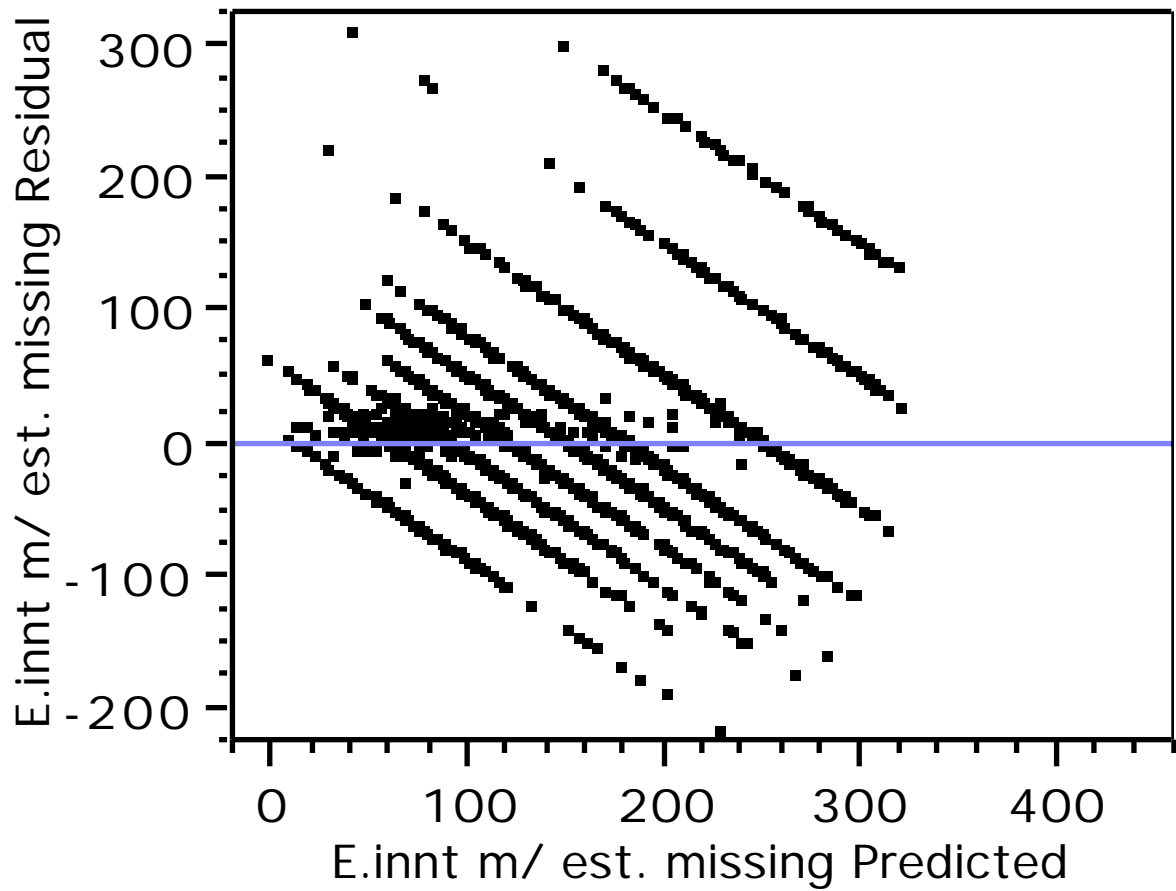
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	22	14147618	643074	210.5868
Error	2924	8929085	3054	<b>Prob &gt; F</b>
C. Total	2946	23076703		0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	-25.07622	12.19911	-2.06	0.0399
Alder	2.7568453	0.490449	5.62	<.0001
Mann	-97.03559	15.51967	-6.25	<.0001
E.utdanning	4.4516546	0.578609	7.69	<.0001
Heiltidsarbeid	47.546914	3.623145	13.12	<.0001
Offentleg sektor	6.9827594	3.466987	2.01	0.0441
Alder*Alder	-0.026278	0.00535	-4.91	<.0001
Alder*Mann	4.5035116	0.689923	6.53	<.0001
Alder*Alder*Mann	-0.040677	0.00757	-5.37	<.0001
E.utdanning*Mann	2.5174164	0.754816	3.34	0.0009
Heiltidsarbeid*Mann	30.471599	5.02033	6.07	<.0001
Offentleg sektor*Mann	-27.21207	5.049417	-5.39	<.0001
Sentrum storby	16.676846	4.037295	4.13	<.0001
Forstad storby	20.907522	3.55477	5.88	<.0001
Småby	9.6634437	3.541895	2.73	0.0064
Tettstad	9.5787324	3.094525	3.10	0.0020
Uoppg bostad	5.4611336	7.536774	0.72	0.4688
Funksjonær	33.986939	3.120362	10.89	<.0001
Sjølvstendig	24.969538	4.828027	5.17	<.0001
Elev/ student	-23.84837	4.194874	-5.69	<.0001
Pensjon/ trygd	-5.282188	4.611503	-1.15	0.2521
Andre KtL	-15.4818	4.084628	-3.79	0.0002
Uoppgitt KtL	-24.72841	8.489929	-2.91	0.0036

**OPPGAVE 2/ OPPGAVE 2 Modell 6 /QUESTION 2 Model 6**

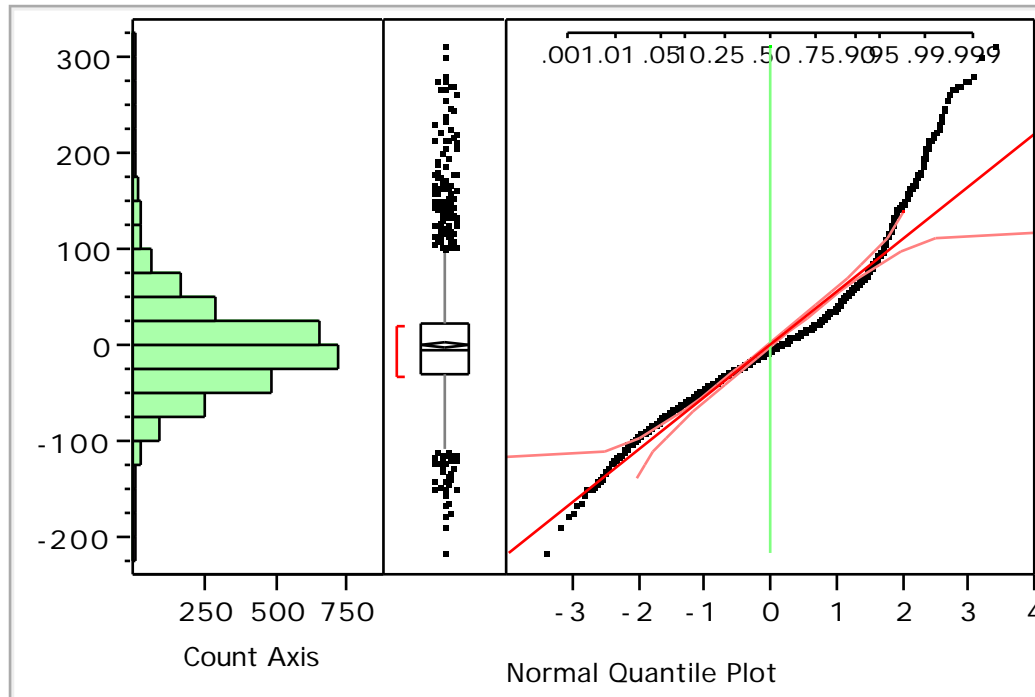
**Residual by Predicted Plot**



**OPPGÅVE 2/ OPPGÅVE 2 Modell 6 /QUESTION 2 Model 6**

**Distributions**

**Residual E.innt(EM) Modell 6**



**Quantiles**

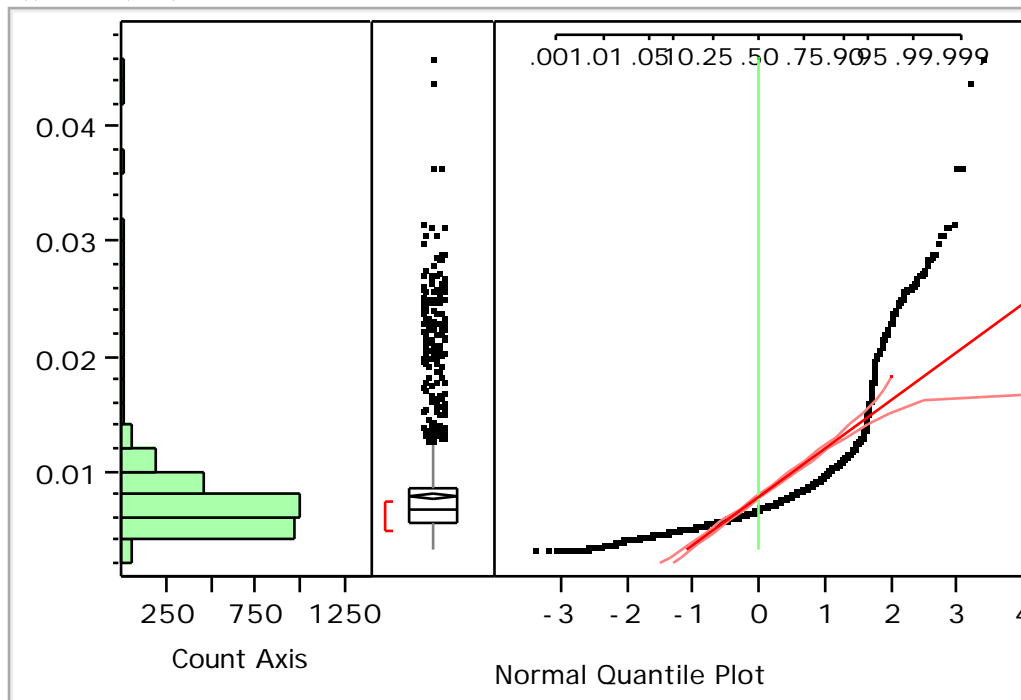
100.0%	maximum	309.88
99.5%		234.28
97.5%		145.69
90.0%		60.53
75.0%	quartile	21.90
50.0%	median	-4.70
25.0%	quartile	-30.85
10.0%		-58.96
2.5%		-91.76
0.5%		-140.37
0.0%	minimum	-217.61

**Moments**

Mean	7.2e-14
Std Dev	55.053775
Std Err Mean	1.0141379
upper 95% Mean	1.9885264
lower 95% Mean	-1.988526
N	2947

**OPPGÅVE/ OPPGAVE 2 Modell 6 /QUESTION 2 Model 6**

**h(i) E.innt(EM) Modell 6**



**Quantiles**

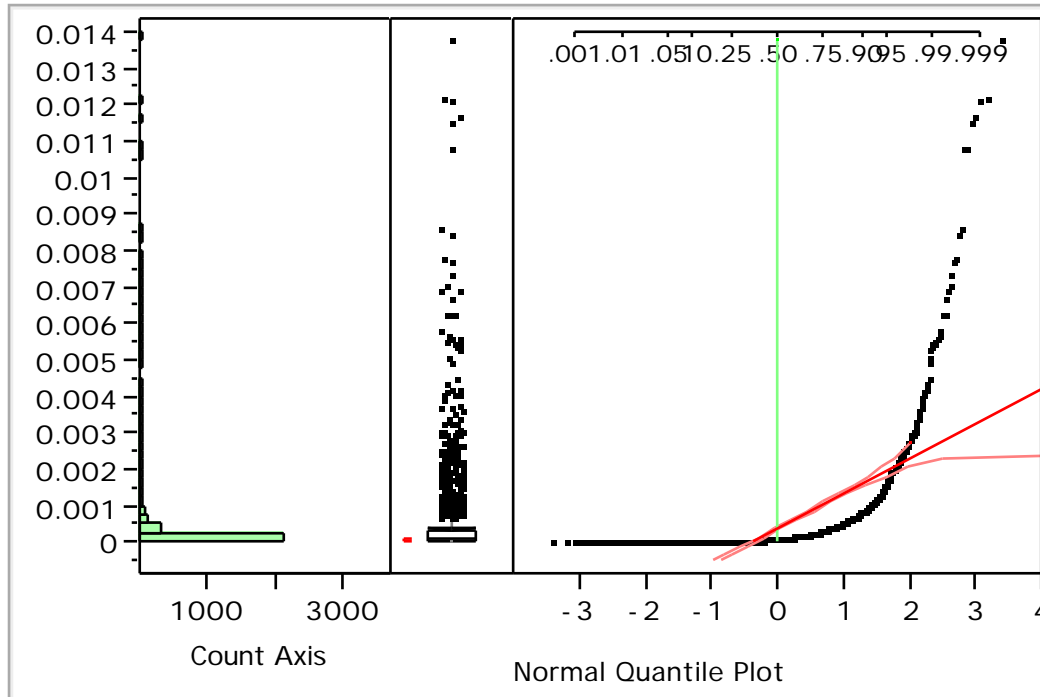
100.0%	maximum	0.04598
99.5%		0.02855
97.5%		0.02262
90.0%		0.01113
75.0%	quartile	0.00844
50.0%	median	0.00671
25.0%	quartile	0.00558
10.0%		0.00481
2.5%		0.00408
0.5%		0.00329
0.0%	minimum	0.00319

**Moments**

Mean	0.0078045
Std Dev	0.0041797
Std Err Mean	0.0000770
upper 95% Mean	0.0079555
lower 95% Mean	0.0076536
N	2947

**OPPGAVE 2/ OPPGAVE 2 Modell 6 /QUESTION 2 Model 6**

**Cook's D(i) E.innt(EM) Modell 6**



**Quantiles**

100.0%	maximum	0.01381
99.5%		0.00687
97.5%		0.00257
90.0%		0.00080
75.0%	quartile	0.00029
50.0%	median	0.00008
25.0%	quartile	0.00001
10.0%		0.00000
2.5%		0.00000
0.5%		0.00000
0.0%	minimum	0.00000

**Moments**

Mean	0.0003548
Std Dev	0.0009597
Std Err Mean	0.0000177
upper 95% Mean	0.0003895
lower 95% Mean	0.0003202
N	2947

**OPPGÅVE 2/ OPPGAVE 2 Modell 7 /QUESTION 2 Model 7****Response E.innt****Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.550587
RSquare Adj	0.548701
Root Mean Square Error	58.39633
Mean of Response	150.7935
Observations (or Sum Wgts)	2634

**Analysis of Variance**

The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

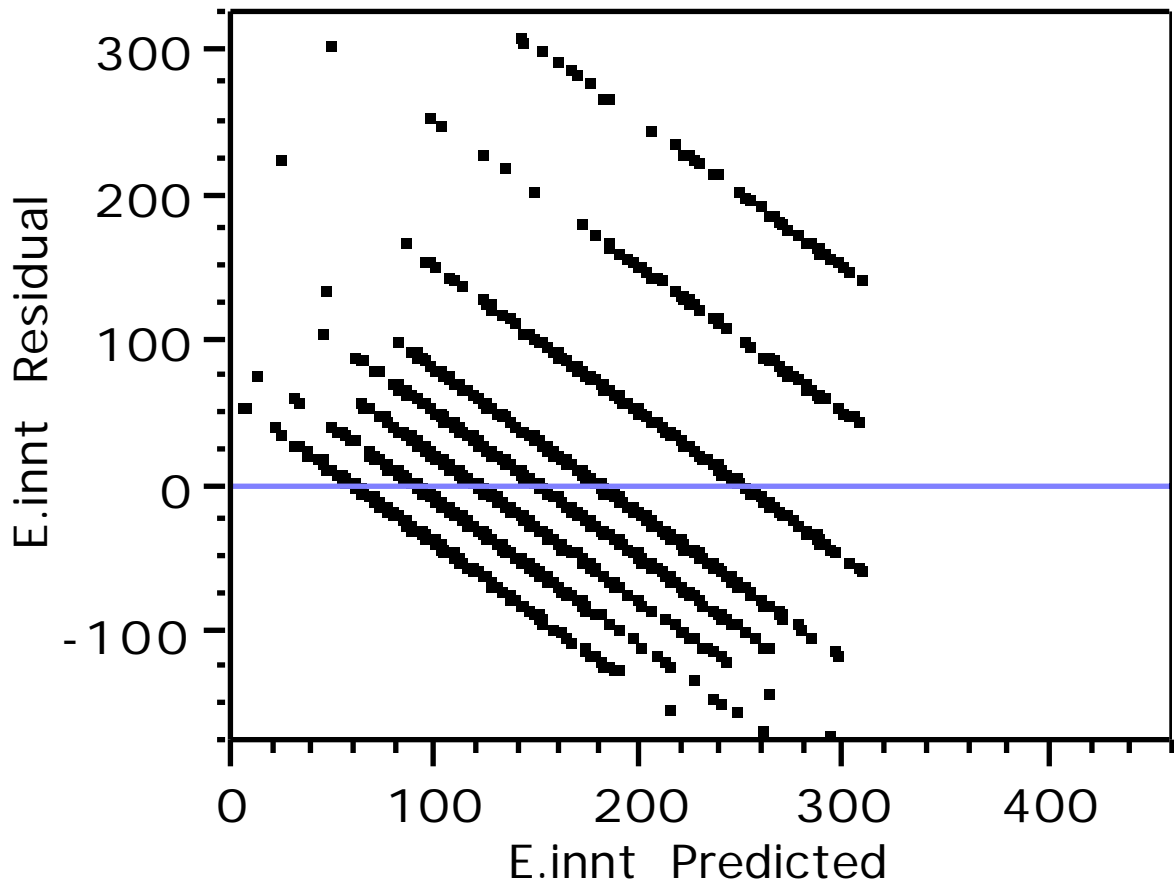
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	11	10954276	995843	292.0249
Error	2622	8941365	3410	<b>Prob &gt; F</b>
C. Total	2633	19895642		0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	-67.00947	12.99689	-5.16	<.0001
Alder	4.8758099	0.532959	9.15	<.0001
Mann	-98.59718	18.00093	-5.48	<.0001
E.utdanning	5.3362693	0.629424	8.48	<.0001
Heiltidsarbeid	62.605666	3.631676	17.24	<.0001
Offentleg sektor	3.5888226	3.651734	0.98	0.3258
Alder*Alder	-0.04714	0.005791	-8.14	<.0001
Alder*Mann	3.9804485	0.78784	5.05	<.0001
Alder*Alder*Mann	-0.034699	0.00859	-4.04	<.0001
E.utdanning*Mann	3.0099967	0.833378	3.61	0.0003
Heiltidsarbeid*Mann	30.257213	5.42762	5.57	<.0001
Offentleg sektor*Mann	-22.35092	5.412688	-4.13	<.0001

**OPPGAVE 2/ OPPGAVE 2 Modell 7 /QUESTION 2 Model 7**

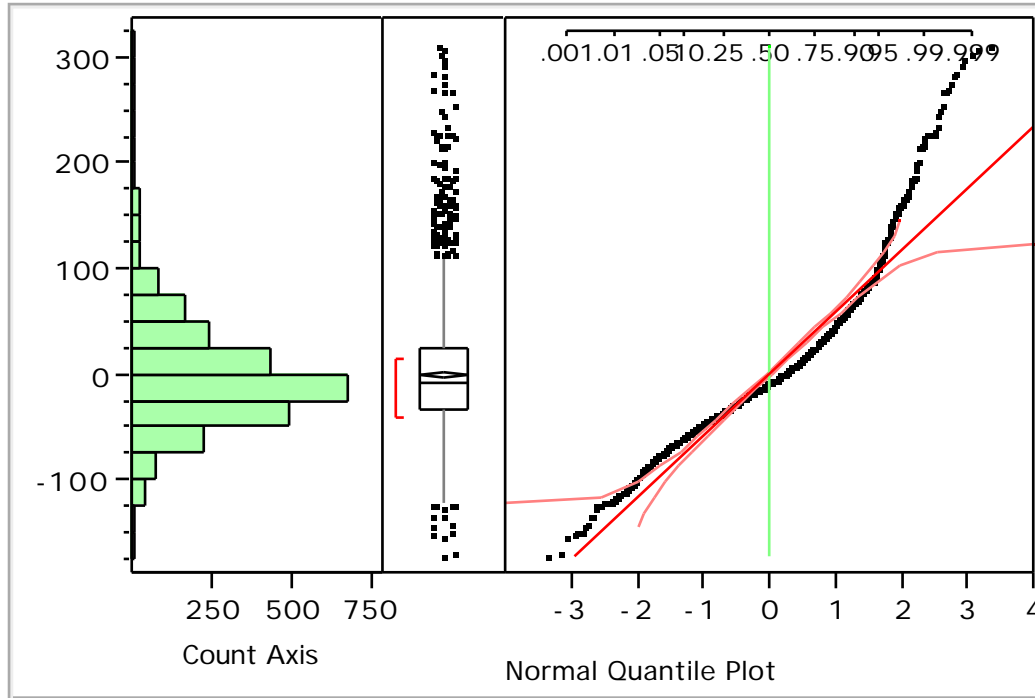
**Residual by Predicted Plot**



**OPPGÅVE 2/ OPPGÅVE 2 Modell 7 /QUESTION 2 Model 7**

**Distributions**

**Residual E.innt Modell 7**



**Quantiles**

100.0%	maximum	308.90
99.5%		248.07
97.5%		157.59
90.0%		68.10
75.0%	quartile	24.09
50.0%	median	-8.01
25.0%	quartile	-34.79
10.0%		-58.87
2.5%		-93.94
0.5%		-124.68
0.0%	minimum	-172.45

**Moments**

Mean	4.3e-15
Std Dev	58.274222
Std Err Mean	1.1354515
upper 95% Mean	2.2265073
lower 95% Mean	-2.226507
N	2634



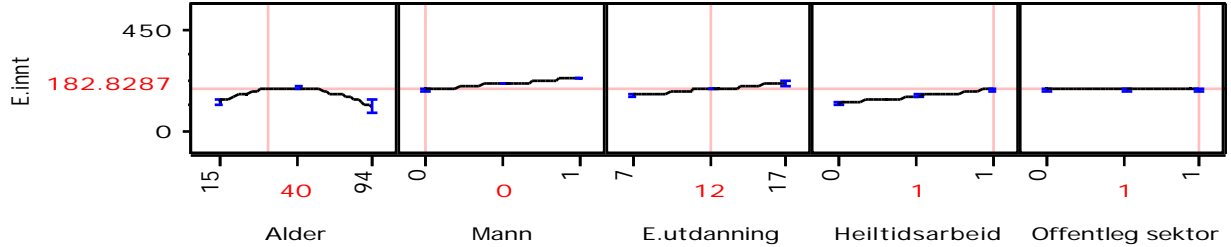
**OPPGAVE 2/ OPPGAVE 2 Modell 7 / QUESTION 2 Model 7**

**Response E.innt**

**Prediction Profile<sup>2</sup>**

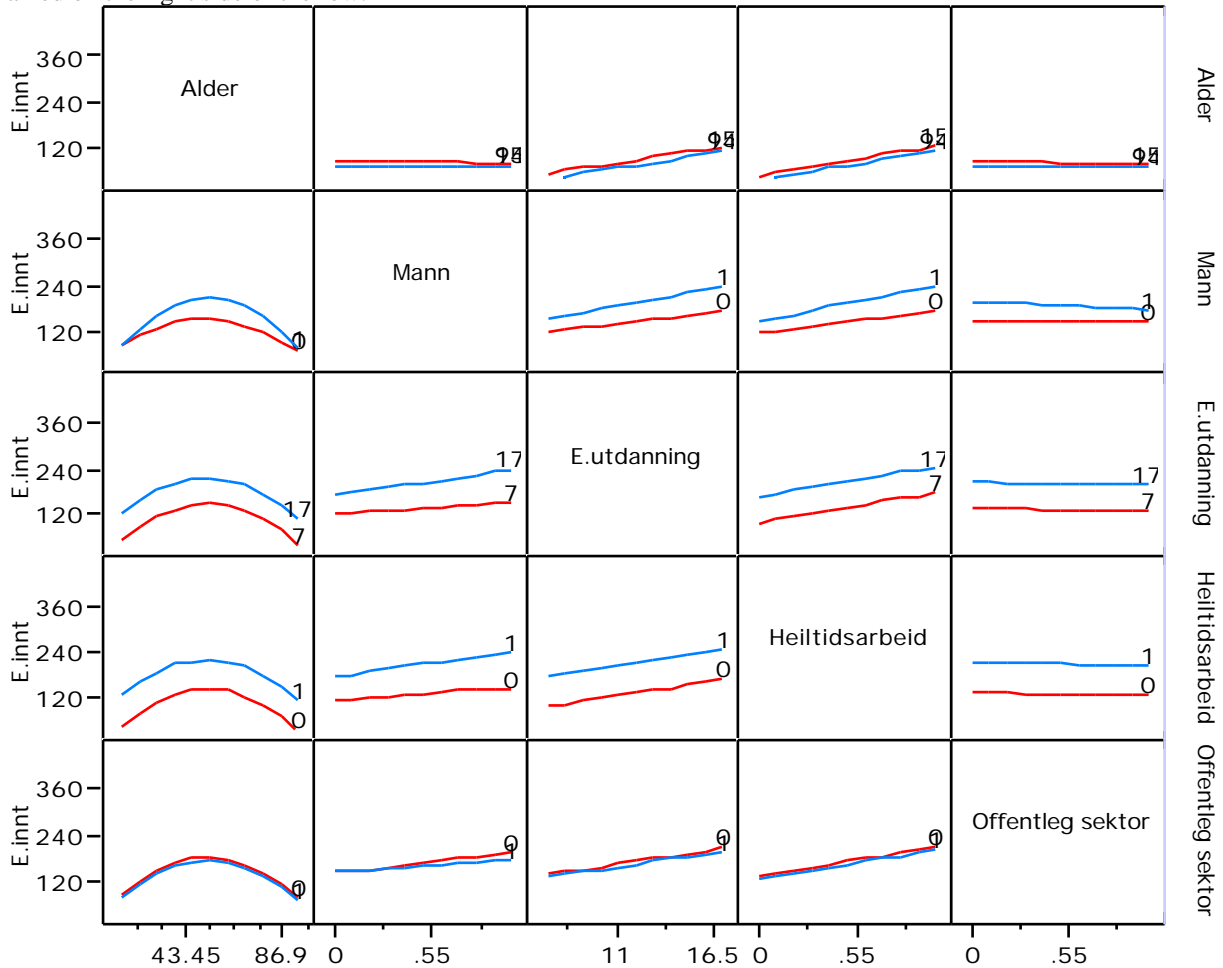
The change in predicted response as you vary one factor at a time, holding the other factors at their current values.

Click in the graph to change the current values of the factors.



**Interaction Profiles**

The change in predicted response by the variable named in the diagonal and for low and high values of the variable named on the right side of the row.



<sup>2</sup> Meir om lesing av plottet er gitt under vedlegg til oppgave 3. Mer om lesing av plottet er gitt under vedlegg til oppgave 3. More on the reading of the plot is presented in attachments to question 3

**OPPGÅVE 2/ OPPGAVE 2 Modell 7 /QUESTION 2 Model 7****Frekvensfordelinger for 314 personer utan inntektsopplysningar****Heiltidsarbeid**

Level	Count	Prob
0	275	0.87580
1	39	0.12420

**Offentleg sektor**

Level	Count	Prob
0	283	0.90127
1	31	0.09873

**E.utdanning**

År	Count	Prob
7	69	0.21975
9	95	0.30255
12	126	0.40127
14	16	0.05096
17	8	0.02548

**Alder 10-Årsgr**

År	Count	Prob
-29	140	0.44586
30-39	56	0.17834
40-49	29	0.09236
50-59	29	0.09236
60-69	18	0.05732
70-79	29	0.09236
80-89	11	0.03503
90+	2	0.00637

**HH.inntekt (1000)**

1000 kr	Count	Prob
60	2	0.00637
90	16	0.05096
120	5	0.01592
150	18	0.05732
180	30	0.09554
250	27	0.08599
350	18	0.05732
450	13	0.04140
998=miss	164	0.52229
999=miss	21	0.06688

---

**OPPGÅVE 2/ OPPGÅVE 2 Modell 7 /QUESTION 2 Model 7****Frekvensfordelinger for 314 personar utan inntektsopplysningar****Busadstype**

Type	Count	Prob
1	17	0.05414
2	43	0.13694
3	74	0.23567
4	105	0.33439
5	62	0.19745
6	13	0.04140

**Kjelde til livsopphald**

Level	Count	Prob
1	31	0.09873
2	25	0.07962
3	7	0.02229
4	3	0.00955
5	4	0.01274
6	94	0.29936
7	22	0.07006
8	27	0.08599
9	24	0.07643
10	41	0.13057
11	10	0.03185
12	26	0.08280

---

**OPPGÅVE 2/ OPPGAVE 2 Modell 8 /QUESTION 2 Model 8****Response E.innt****Summary of Fit**

R-square is the portion of variation attributed to the model, between 0 and 1. Root Mean Squared Error "RMSE" estimates the standard deviation of the residual.

RSquare	0.587268
RSquare Adj	0.58379
Root Mean Square Error	56.08024
Mean of Response	150.7935
Observations (or Sum Wgts)	2634

**Analysis of Variance**

The test that the whole model fits better than a simple mean, i.e. testing that all the parameters are zero except the intercept

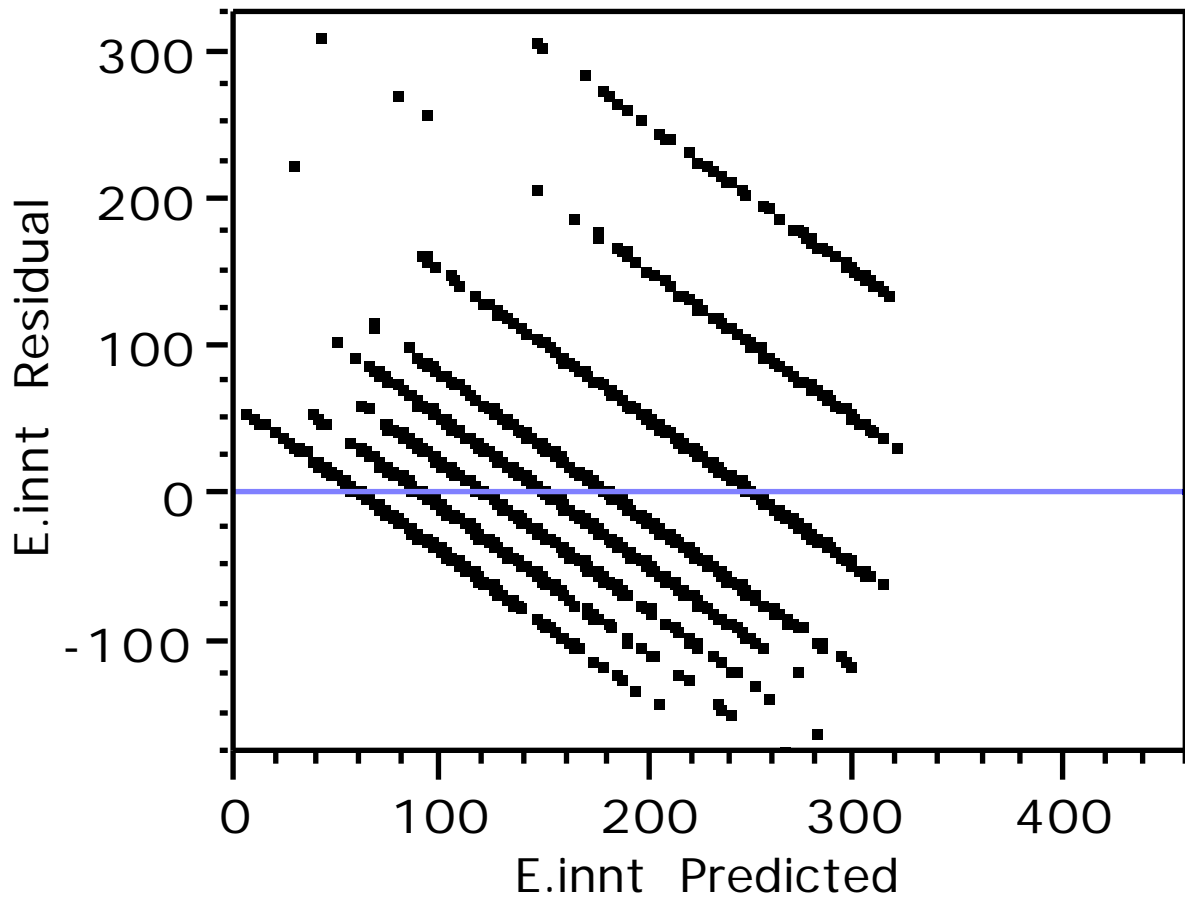
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	22	11684064	531094	168.8696
Error	2611	8211578	3145	<b>Prob &gt; F</b>
C. Total	2633	19895642		0.0000

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	-20.08853	13.60022	-1.48	0.1398
Alder	2.8609311	0.545856	5.24	<.0001
Mann	-93.22601	17.35376	-5.37	<.0001
E.utdanning	4.3121929	0.6226	6.93	<.0001
Heiltidsarbeid	44.612115	3.77656	11.81	<.0001
Offentleg sektor	2.7964771	3.648578	0.77	0.4435
Alder*Alder	-0.027309	0.005919	-4.61	<.0001
Alder*Mann	4.3286729	0.760616	5.69	<.0001
Alder*Alder*Mann	-0.038914	0.008295	-4.69	<.0001
E.utdanning*Mann	2.3878568	0.803484	2.97	0.0030
Heiltidsarbeid*Mann	28.749835	5.324208	5.40	<.0001
Offentleg sektor*Mann	-19.90565	5.261246	-3.78	0.0002
Sentrum storby	14.355652	4.276097	3.36	0.0008
Forstad storby	20.290997	3.819402	5.31	<.0001
Småby	11.641376	3.862122	3.01	0.0026
Tettstad	9.5568266	3.347434	2.85	0.0043
Uoppb bostad	8.7825941	8.567582	1.03	0.3054
Funksjonær	34.188963	3.231401	10.58	<.0001
Sjølvstendig	25.484098	4.99213	5.10	<.0001
Elev/ student	-26.85123	4.532515	-5.92	<.0001
Pensjon/ trygd	-9.319354	4.919899	-1.89	0.0583
Andre KtL	-6.313265	4.591924	-1.37	0.1693
Uoppgitt KtL	10.739994	12.04012	0.89	0.3725

**OPPGAVE 2/ OPPGAVE 2 Modell 8 /QUESTION 2 Model 8**

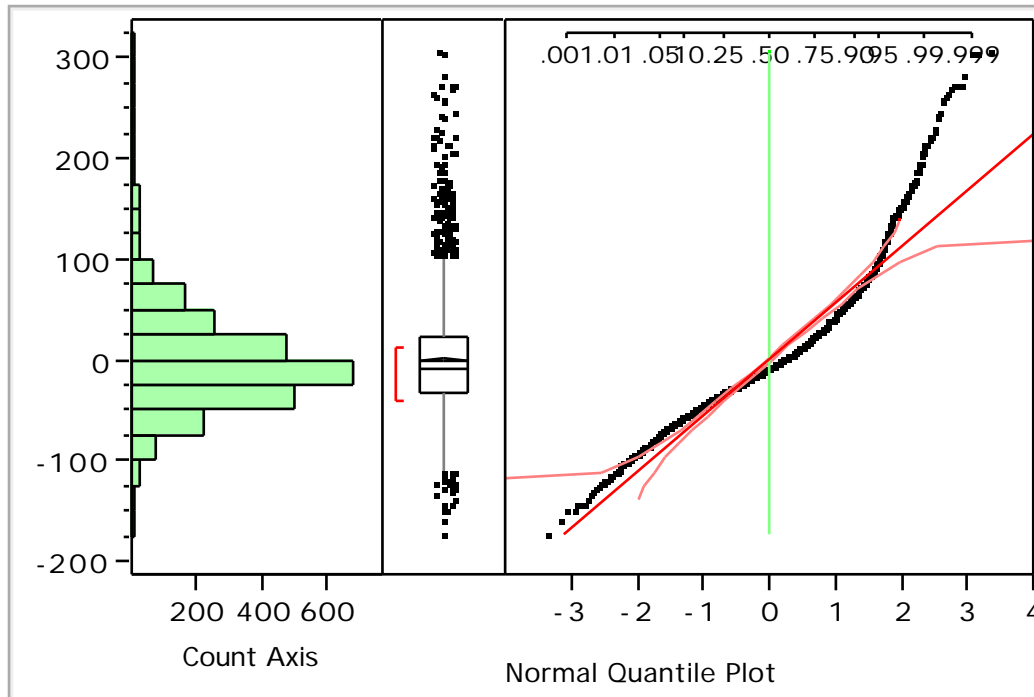
**Residual by Predicted Plot**



**OPPGÅVE 2/ OPPGÅVE 2 Modell 8 /QUESTION 2 Model 8**

**Distributions**

**Residual E.innt Modell 8**



**Quantiles**

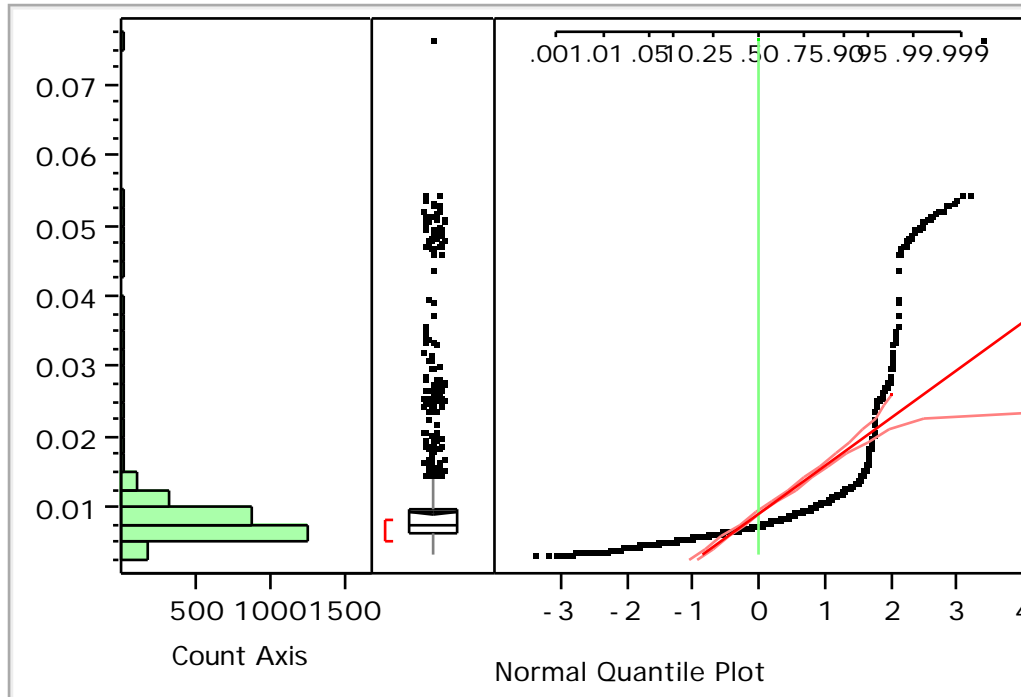
100.0%	maximum	307.61
99.5%		244.85
97.5%		150.01
90.0%		63.94
75.0%	quartile	22.26
50.0%	median	-8.17
25.0%	quartile	-31.72
10.0%		-56.86
2.5%		-90.95
0.5%		-125.81
0.0%	minimum	-174.37

**Moments**

Mean	0.000000
Std Dev	55.845461
Std Err Mean	1.0881280
upper 95% Mean	2.1337106
lower 95% Mean	-2.133711
N	2634

**OPPGAVE 2/ OPPGAVE 2 Modell 8 /QUESTION 2 Model 8**

h(i) E.innt Modell 8



**Quantiles**

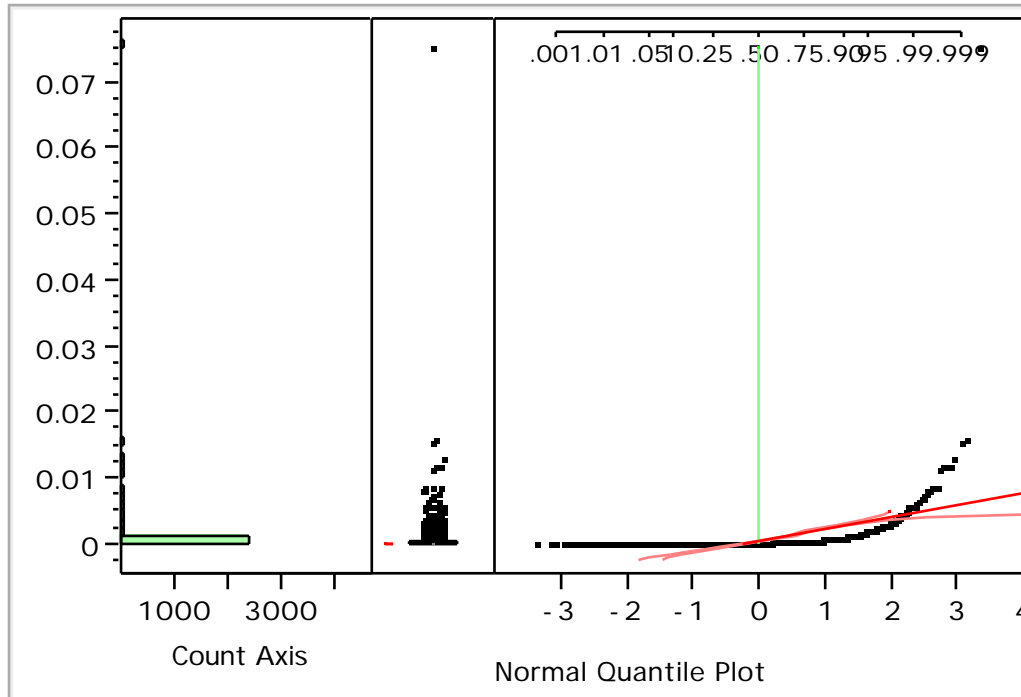
100.0%	maximum	0.07654
99.5%		0.05118
97.5%		0.02779
90.0%		0.01238
75.0%	quartile	0.00958
50.0%	median	0.00755
25.0%	quartile	0.00627
10.0%		0.00530
2.5%		0.00443
0.5%		0.00351
0.0%	minimum	0.00340

**Moments**

Mean	0.0091540
Std Dev	0.0067386
Std Err Mean	0.0001241
upper 95% Mean	0.0093974
lower 95% Mean	0.0089107
N	2948

**OPPGAVE 2/ OPPGAVE 2 Modell 8 /QUESTION 2 Model 8**

**Cook's D(i) E.innt Modell 8**



**Quantiles**

100.0%	maximum	0.07516
99.5%		0.00810
97.5%		0.00285
90.0%		0.00086
75.0%	quartile	0.00030
50.0%	median	0.00009
25.0%	quartile	0.00002
10.0%		0.00000
2.5%		0.00000
0.5%		0.00000
0.0%	minimum	0.00000

**Moments**

Mean	0.0004171
Std Dev	0.0018016
Std Err Mean	0.0000351
upper 95% Mean	0.0004859
lower 95% Mean	0.0003483
N	2634



**OPPGÅVE 3 Variabeldefinisjoner/ OPPGAVE 3 Variabeldefinisjoner/  
QUESTION 3 Definition of variables**

**Besøke lokalt kunstgalleri** = 1 dersom «Val av tidsbruk i bygdenorge» = 1, 0 elles

Val av tidsbruk i bygdenorge er svar på spørsmålet

Dersom du var på reise i bygdenorge og måtte vente ein time på ferje, kva for ein av desse stadane ville du helst vitje? Lokalt kusntgalleri, husflidsforretning, staden sitt kjøpsenter, kafe/ restaurant.

**Besøke lokalt kunstgalleri** = 1 hvis «Val av tidsbruk i bygdenorge» = 1, 0 ellers

Val av tidsbruk i bygdenorge er svar på spørsmålet

«Hvis du var på reise i bygdenorge og måtte vente en time på ferge, hvilke av disse stende ville du da helst besøke? Lokalt kusntgalleri, husflidsforretning, stedets kjøpsenter, kafe/ restaurant.»

**Besøke lokalt kunstgalleri** (Visit local art gallery ) = 1 if «Val av tidsbruk i bygdenorge» (choice of spending time in rural Norway) = 1, 0 otherwise

Val av tidsbruk i bygdenorge is the answer to the question

If you were travelling in Norway and had to wait an hour for the ferry, which of the following establishments would you rather visit? Local art gallery, arts and crafts shop, local shopping mall, café/ restaurant.

Frequencies

Level	Count	Prob
1	393	0.13331
0	2555	0.86669
Total	2948	1.00000

**Kvinne**

**Woman**

Frequencies

Level	Count	Prob
0	1445	0.49016
1	1503	0.50984
Total	2948	1.00000

**E.utdanning**

sjå oppgåve 2 variablar

**E.utdanning**

se oppgave 2 variabler

**Own education**

see question 2 variables

**Busadstype**

sjå oppgåve 2 variablar

**Bostedstype**

se oppgave 2 variabler

**Type of home town**

see question 2 variables

**Alder** = alder i år

sjå oppgåve 2 variablar

**Alder** = alder i år

se oppgave 2 variabler

**Alder** (age) = age in years

see question 2 variables

## OPPGÅVE 3 Variabeldefinisjoner/ OPPGAVE 3 Variabeldefinisjoner/ QUESTION 3 Definition of variables

**Mors utdanning** = «mor si utdanning i år» er koda

7 dersom mor har utdanning på folkeskolenivå (inntil 8 års skolegang)

9 dersom mor har utdanning på ungdomsskole/ realskolenivå (9-10 års skolegang)

12 dersom mor har utdanning på videregående skole/ gymnas nivå (11-13 års skolegang)

15 dersom mor har utdanning på universitetsnivå (meir enn 12 års skolegang + studiar),

**Mors utdanning** = «mors utdanning i år» er kodet

7 hvis mor har utdanning på folkeskolenivå (inntil 8 års skolegang)

9 hvis mor har utdanning på ungdomsskole/ realskolenivå (9-10 års skolegang)

12 hvis mor har utdanning på videregående skole/ gymnas nivå (11-13 års skolegang)

15 hvis mor har utdanning på universitetsnivå (mer enn 12 års skolegang + studier),

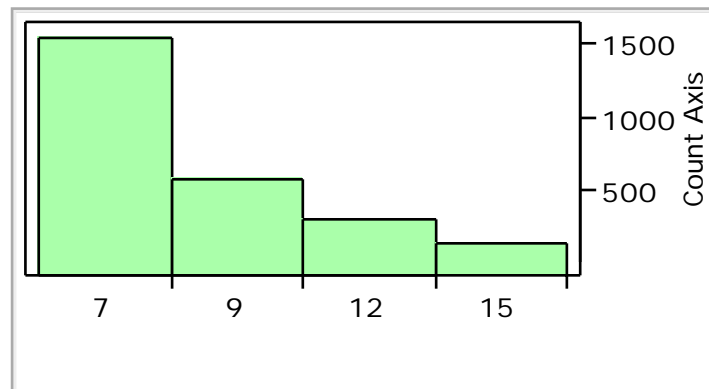
**Mors utdanning** = «mother's education in years» has been coded

7 if mother has education at the level of primary school (up to 8 years of education)

9 if mother has education at the level of secondary school (9-10 years of education)

12 if mother has education at the level of high school (11-13 years of education)

15 if mother has education at college level (more than 12 years of education + college studies),



### Frequencies

Level	Count	Prob
7	1654	0.56106
9	663	0.22490
12	402	0.13636
15	229	0.07768
Total	2948	1.00000

**OPPGÅVE 3/ OPPGAVE 3 Modell 1-4 /  
QUESTION 3 Model 1-4****Kommentarar til vedleggstabellane for oppgåve 3****1. Oddsrate**

Til skilnad frå SPSS vil kolonna med oddsrate her gi høvestalet mellom oddsen for å ha variabelen sin høgaste verdi og oddsen for å ha variabelen sin lågaste verdi.

**2. Rsquare**

Rsquare(U) for ein modell er definert som reduksjonen i negativ loglikelihood som følgjer av modellen (kalla Difference, skilnaden mellom Full og Reduced) i høve til den negative loglikelihooden i ein modell som berre har konstant til parameter (kalla Reduced)

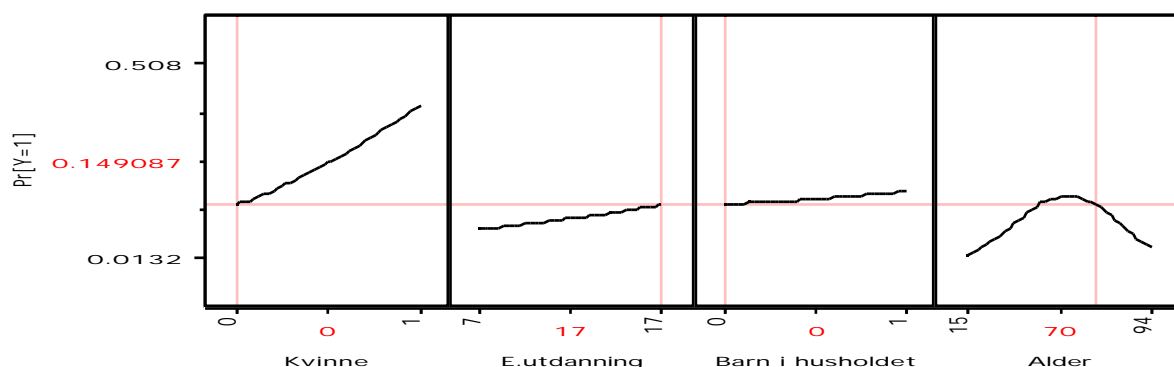
**3. VIF**

VIF = variansinflasjonsfaktoren =  $1/\text{toleransen}$

**4. Betinga effekt plott**

I ei rekkje av betinga effekt plott (eksempel nedanfor) vil kvart einskild plott gi oss samanhengen mellom sannsynet for den gitte avhengige variabelen ( $\Pr\{Y=1\}$ ) og den uavhengige som er nemnt under det einskilde plottet når dei andre uavhengige variablane er haldne konstante med dei verdiane som står midt under plottet deira. Variabelverdiane som står på kvar side gir minste og største observerte variabelverdi for variabelen. Det samme gjeld for verdiane som er gitt for sannsynet.

Eksempel: Samanhengen mellom «Kvinne» og  $\Pr\{Y=1\}$  er betinga av variabelverdiane Eg.Utd. = 17, Barn i HH = 0 og Alder = 70. Største moglege verdi av  $\Pr\{Y=1\} = 0.508$  og lågaste moglege verdi er 0.0132. For Kvinne = 0 (dvs. for ein mann), gitt at dei andre variablane er faste (Eg.Utd. = 17, Barn i HH = 0 og Alder = 70), vil  $\Pr\{Y=1\} = 0.149087$ , talet som er gitt midt på skalaen.



**OPPGÅVE 2/ OPPGAVE 3 Modell 1-4 /  
QUESTION 3 Model 1-4****Kommentarer til vedleggstabellene for oppgave 3****1. Oddsrat**

Ulikt SPSS vil kolonnen med oddsrat her gi forholdstallet mellom oddsen for å ha variabelens høyeste verdi og oddsen for å ha variabelens laveste verdi.

**2. Rsquare**

Rsquare(U) for en modell er definert som reduksjonen i negativ loglikelihood som følger av modellen (kalt Difference, forskjellen mellom Full og Reduced) i forhold til den negative loglikelihooden i en modell som bare har konstant til parameter (kalt Reduced)

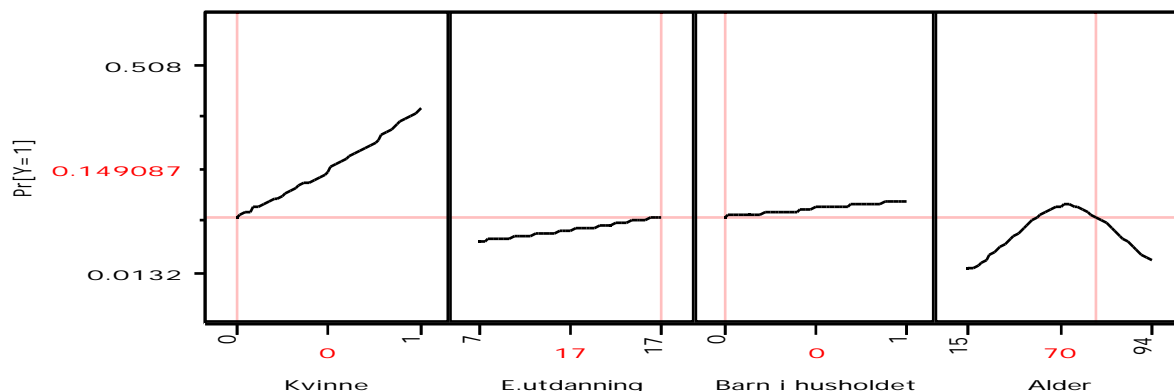
**3. VIF**

VIF = variansinflasjonsfaktoren =  $1 / \text{toleransen}$

**4. Betinget effekt plott**

I en rekke av betinget effekt plott (eksempel nedenfor) vil hvert enkelt plott gi oss sammenhengen mellom sannsynligheten for den gitte avhengige variabelen ( $\Pr\{Y=1\}$ ) og den uavhengige som er nevnt under det enkelte plottet når de andre uavhengige variablene er holdt konstante med de verdiene som står midt under plottet deres. Variabelverdiene som står på hver side gir minste og største observerte variabelverdi for variabelen. Det samme gjelder for verdiene som er gitt for sannsynligheten.

Eksempel: Sammenhengen mellom «Kvinne» og  $\Pr\{Y=1\}$  er betinget av variabelverdiene Eg.Utd. = 17, Barn i HH = 0 og Alder = 70. Største mulige verdi av  $\Pr\{Y=1\} = 0.508$  og laveste mulige verdi er 0.0132. For Kvinne = 0 (dvs. for en mann), gitt at de andre variablene er faste (Eg.Utd. = 17, Barn i HH = 0 og Alder = 70), vil  $\Pr\{Y=1\} = 0.149087$ , tallet som er gitt midt på skalaen.



## OPPGÅVE 3/ OPPGAVE 3 Modell 1-4 / QUESTION 3 Model 1-4

### Comments to the tables attached to question 3

#### 1. Odds ratios

Different from SPSS the column of odds ratios here gives the ratio of the odds of having the highest value of the variable and the odds of having the lowest value of the variable.

#### 2. Rsquare

Rsquare(U) for a model is defined as the reduction in i negativ loglikelihood caused by the model (called Difference, the difference between Full and Reduced) in relation to the negative loglikelihood in a model with only a constant as parameter (called Reduced)

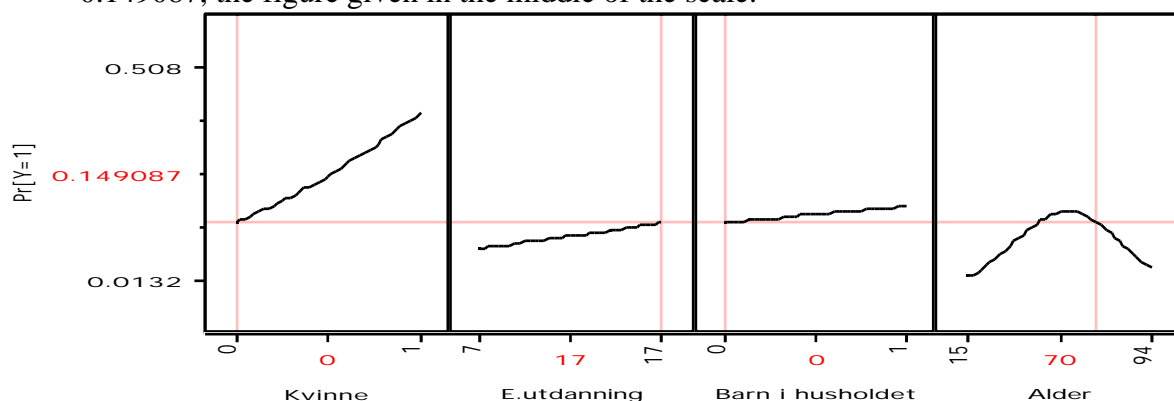
#### 3. VIF

VIF = Variance inflation factor =  $1 / \text{tolerance}$

#### 4. Conditional effect plots

In a row of conditional effect plots (as in the example below) each plot will give us the relationship between the probability of the dependent variable (  $\Pr\{Y=1\}$  ) and the independent variable listed below each plots with the other independent variables kept konstant at the values given in the middle below their plot. The variable values on each side give the lowest and highest observed value on that variable. Similarly applies for the values given for the probability.

Example: the relationship between «Kvinne» and  $\Pr\{Y=1\}$  is based on the variable values Eg.Utd. = 17, Barn i HH = 0 and Alder = 70. The highest possible value of  $\Pr\{Y=1\}= 0.508$  and the lowest possible value is 0.0132. For Kvinne = 0 (i.e. for a man), given that the other variable values are fixed (i.e. Eg.Utd. = 17, Barn i HH = 0 og Alder = 70),  $\Pr\{Y=1\}= 0.149087$ , the figure given in the middle of the scale.



**OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3****Nominal Logistic Fit for Besøke lokalt kunstgalleri (Modell 1)****Iteration History**

Iter	LogLikelihood	Step	Delta-Criterion	Obj-Criterion
1	-2043.397888	Initial	3334461385	.
2	-1139.051608	Newton	0.94118446	0.79393974
3	-1086.093888	Newton	0.21213971	0.04875935
4	-1083.456556	Newton	0.01574722	0.00243416
5	-1083.444345	Newton	0.00009116	0.00001127

**Whole Model Test**

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	74.0352	4	148.0704	<.0001
Full	1083.4443			
Reduced	1157.4795			

RSquare (U)	0.0640
Observations (or Sum Wgts)	2948

**Parameter Estimates**

Term	Estimate	Std Error	ChiSquare	Prob > ChiSq	Odds Ratio	VIF
Intercept	-5.806519	0.376822	237.44	<.0001	.	.
Kvinne	0.26946785	0.1126913	5.72	0.0168	1.30926754	1.0042269
E.utdanning	0.18127617	0.0195032	86.39	<.0001	6.12734606	1.1762030
Mors utdanning	0.11441208	0.0225642	25.71	<.0001	2.4975161	1.2684774
Alder	0.01530247	0.0036382	17.69	<.0001	3.34978269	1.2455366

**OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3****Nominal Logistic Fit for Besøke lokalt kunstgalleri (Modell 2)****Iteration History**

Iter	LogLikelihood	Step	Delta-Criterion	Obj-Criterion
1	-2043.397888	Initial	2836182339	.
2	-1137.755567	Newton	1.2552844	0.79598324
3	-1084.869293	Newton	0.24279856	0.04874853
4	-1082.237991	Newton	0.03176201	0.00243133
5	-1082.225247	Newton	0.00057358	0.00001178
6	-1082.225247	Newton	0.00000004	3.77e-10

**Whole Model Test**

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	75.2543	5	150.5086	<.0001
Full	1082.2252			
Reduced	1157.4795			

RSquare (U)	0.0650
Observations (or Sum Wgts)	2948

**Parameter Estimates**

Term	Estimate	Std Error	ChiSquare	Prob > ChiSq	Odds Ratio	VIF
Intercept	-4.9510165	0.6624186	55.86	<.0001	.	.
Kvinne	0.2785316	0.1129955	6.08	0.0137	1.32118836	1.00572
E.utdanning	0.10470375	0.0527096	3.95	0.0470	2.84919779	8.41106
Mors utdanning	0.11825219	0.0226564	27.24	<.0001	2.57543251	1.287303
Alder	-0.0045234	0.0133051	0.12	0.7339	0.69952867	15.535473
E.utdanning*Alder	0.00175226	0.0011262	2.42	0.1197	10.7813381	16.498692

**OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3****Nominal Logistic Fit for Besøke lokalt kunstgalleri (Modell 3)****Iteration History**

Iter	LogLikelihood	Step	Delta-Criterion	Obj-Criterion
1	-2043.397888	Initial	1311115136	.
2	-1132.942202	Newton	0.94140722	0.8036135
3	-1079.331806	Newton	0.21611145	0.04966953
4	-1076.75199	Newton	0.0163653	0.0023959
5	-1076.740051	Newton	0.00009596	0.00001109

**Whole Model Test**

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	80.7395	7	161.479	<.0001
Full	1076.7401			
Reduced	1157.4795			

RSquare (U)	0.0698
Observations (or Sum Wgts)	2948

**Parameter Estimates**

Term	Estimate	Std Error	Chi-Square	Prob > ChiSq	Odds Ratio	VIF
Intercept	-2.2630382	1.6045341	1.99	0.1584	.	.
Kvinne	0.26434937	0.1133831	5.44	0.0197	1.3025832	1.00762
E.utdanning	-0.2060165	0.1357997	2.30	0.1293	0.12743297	52.67944
Mors utdanning	0.13107263	0.0233755	31.44	<.0001	2.85359914	1.34854
Alder	-0.1325838	0.0721679	3.38	0.0662	0.00002826	450.71094
E.utdanning*Alder	0.01652043	0.0061712	7.17	0.0074	5446439797	456.86846
Alder*Alder	0.00135636	0.000742	3.34	0.0676	118135.771	409.98939
Alder*Alder*E.utdanning	-0.0001597	0.0000646	6.10	0.0135	2.462e-9	352.68711



**OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3**

**Tabell "MostLikely Besøke lokalt kunstgalleri" mot**

**"Besøke lokalt kunstgalleri" (Modell 3)**

Count	Most likely	0	
Total %	Y=1		
Col %			
Row %			
<b>Observed</b>	4	389	393
<b>Y=1</b>	0.14	13.20	13.33
	26.67	13.26	
	1.02	98.98	
<b>0</b>	11	2544	2555
	0.37	86.30	86.67
	73.33	86.74	
	0.43	99.57	
	15	2933	2948
	0.51	99.49	

**Tests**

Source	DF	-LogLike	RSquare (U)
Model	1	0.941298	0.0100
Error	2946	93.232957	
C. Total	2947	94.174254	
N	2948		

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	1.883	0.1700
Pearson	2.321	0.1277

**Fisher's Exact Test**

	Prob
Left	0.9606
Right	0.1286
2-Tail	0.1286

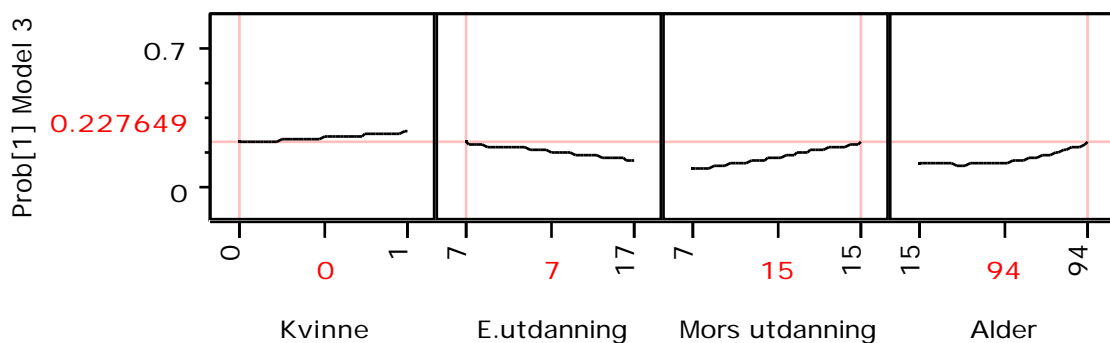
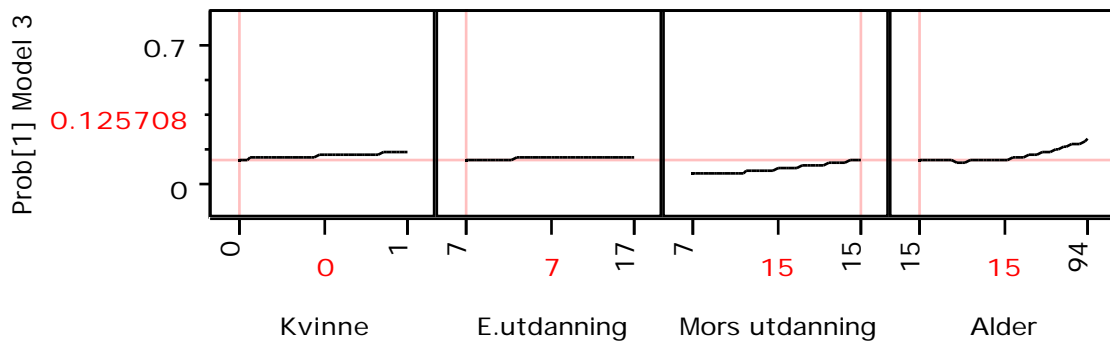
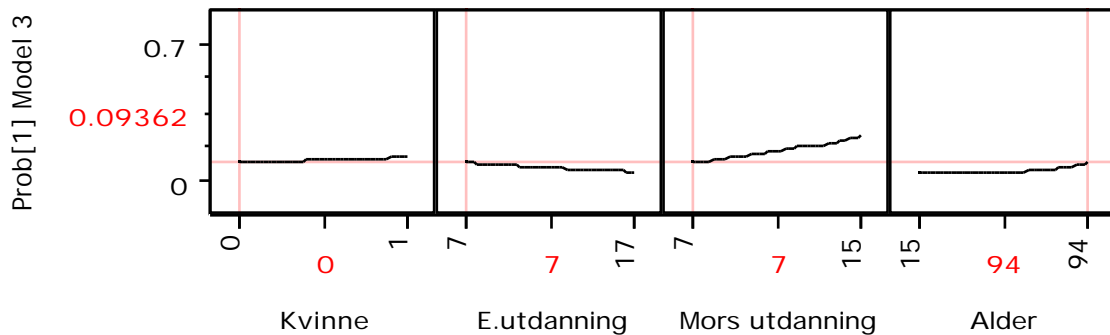
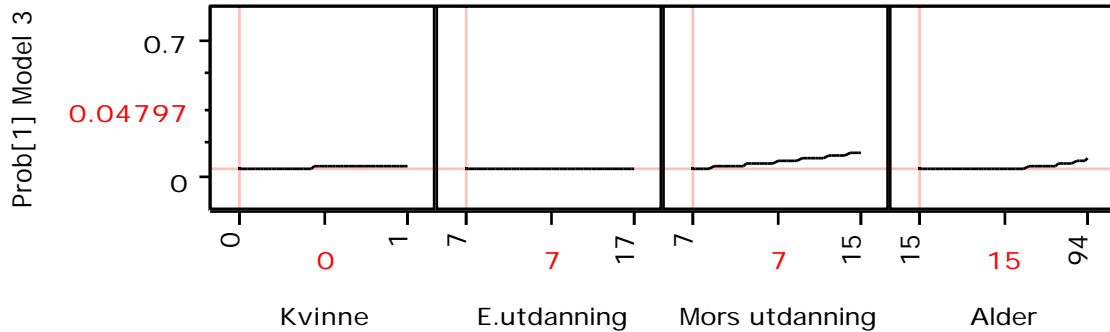
Kappa	Std Err
0.009903	0.008793

Kappa measures the degree of agreement.

### OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3

#### Betinga effekt plott (Modell 3)

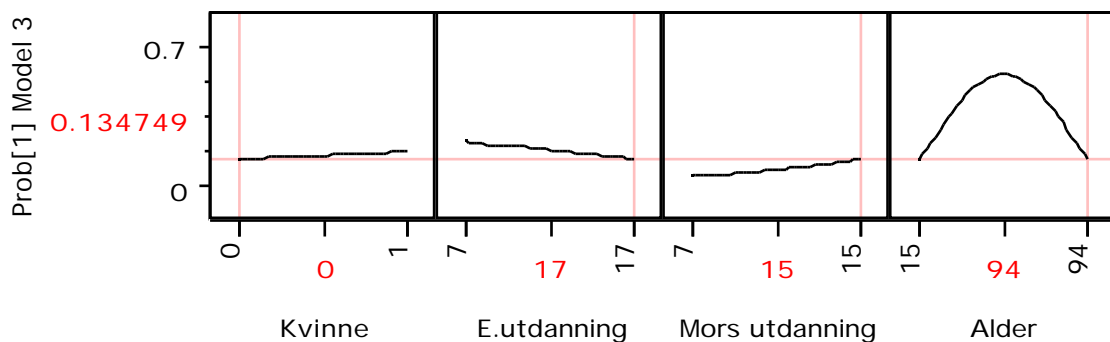
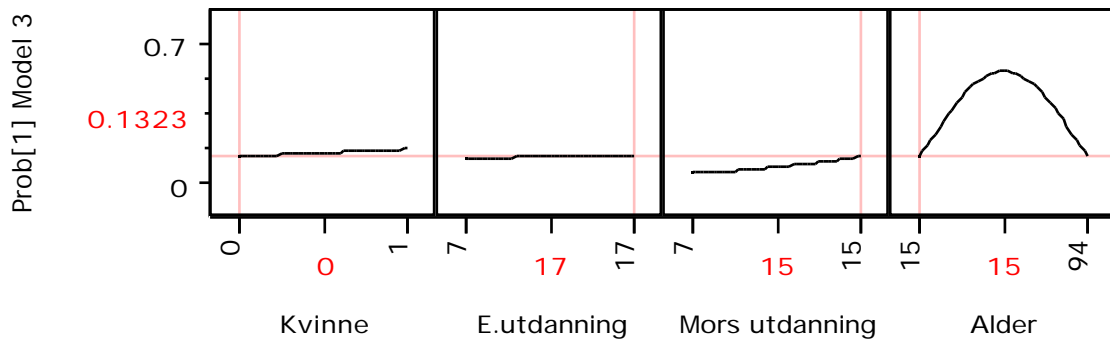
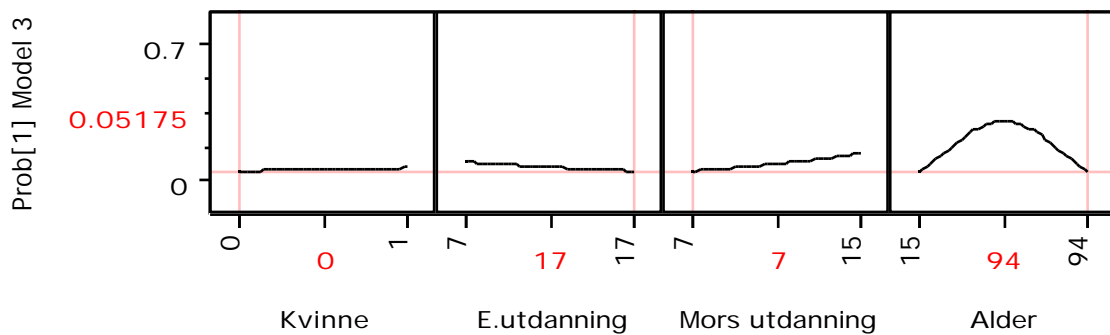
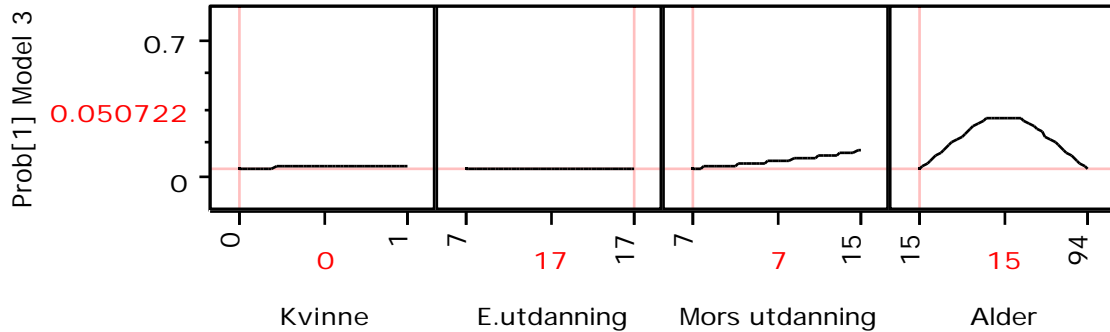
The change in predicted response as you vary one factor at a time, holding the other factors at their current values. Click in the graph to change the current values of the factors.



### OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3

#### Betinga effekt plott (Modell 3)

The change in predicted response as you vary one factor at a time, holding the other factors at their current values. Click in the graph to change the current values of the factors.



**OPPGÅVE 3 /OPPGAVE 3 /QUESTION 3****Nominal Logistic Fit for Besøke lokalt kunstgalleri (Modell 4)****Iteration History**

Iter	LogLikelihood	Step	Delta-Criterion	Obj-Criterion
1	-2043.397888	Initial	1214106150	.
2	-1126.70112	Newton	1.68038031	0.81360408
3	-1069.891802	Newton	0.54284021	0.05309769
4	-1066.795303	Newton	0.06746997	0.00290259
5	-1066.775824	Newton	0.00093657	0.00001826
6	-1066.775823	Newton	0.00000011	1.15e-9

**Whole Model Test**

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	90.7037	12	181.4074	<.0001
Full	1066.7758			
Reduced	1157.4795			

RSquare (U) 0.0784  
 Observations (or Sum Wgts) 2948

**Parameter Estimates**

Term	Estimate	Std Error	Chi-Square	Prob > ChiSq	Odds Ratio	VIF
Intercept	-2.0989542	1.6080969	1.70	0.1918	.	.
Kvinne	0.26397597	0.114002	5.36	0.0206	1.3020969	1.01077
E.utdanning	-0.247483	0.1356535	3.33	0.0681	0.08417733	52.94834
Mors utdanning	0.11756567	0.0236847	24.64	<.0001	2.56132668	1.37755
Alder	-0.1388798	0.0721835	3.70	0.0544	0.00001718	451.87410
E.utdanning*Alder	0.0173218	0.006172	7.88	0.0050	1.616e+10	457.66076
Alder*Alder	0.00140234	0.0007437	3.56	0.0593	175524.169	411.15344
Alder*Alder*E.utdanning	-0.0001657	0.0000647	6.55	0.0105	1.168e-9	353.24315
Sentrum storby	0.75375453	0.2237962	11.34	0.0008	2.1249633	1.64720
Forstad storby	0.56692945	0.2086005	7.39	0.0066	1.76284582	1.85049
Småby	0.52327448	0.214716	5.94	0.0148	1.68754444	1.70990
Tettstad	0.13174341	0.2031617	0.42	0.5167	1.14081556	2.01696
Uoppg bostad	0.18032952	0.4695764	0.15	0.7010	1.19761194	1.10733