MANCHESTER 1824
The University of Manchester

Social Statistics

The Cathie Marsh Institute for Social Research

Longitudinal data across life: *an introduction to cohort data, and its uses in social and health research*

Thursday January 26th 2017

"***Multilevel modelling approach to analysing socioeconomic status longitudinal data and compensating for missingness***"

Email: adrian.byrne@manchester.ac.uk

Talk outline:

1. NCDS survey response

2. Life course response variable

3. Missing data in relation to the response variable

4. Life course multilevel model (MLM)

5. Fixed and random effect results

6. Concluding remarks

NCDS Survey Response

Sweep	PMS	NCDS1	NCDS2	NCDS3	NCDS4	NCDS5	NCDS6	Bio-medical	NCDS7	NCDS8	NCDS9
Year	1958	1965	1969	1974	1981	1991	2000	2002	2004	2008	2013
Age	Birth	7	11	16	23	33	42	44	46	50	55
Productive	17415	15425	15337	14654	12537	11469	11419	9377	9534	9790	9137
Refusal	0	80	797	1151	915	1365	1148	2829	1448	1214	582
Non-contact	218	1036	406	786	1675	1394	1832	792	612	835	860
Other unproductive	0	173	202	295	413	953	263	31	109	332	491
Ineligible	0	0	0	0	0	0	13	65	11	81	0
Not issued *	925	548	275	0	862	993	1415	2908	4248	3553	4543
Emigrant	0	475	701	799	1196	1335	1268	1234	1272	1293	1286
Dead	0	821	840	873	960	1049	1200	1322	1324	1460	1659
Total	18558	18558	18558	18558	18558	18558	18558	18558	18558	18558	18558

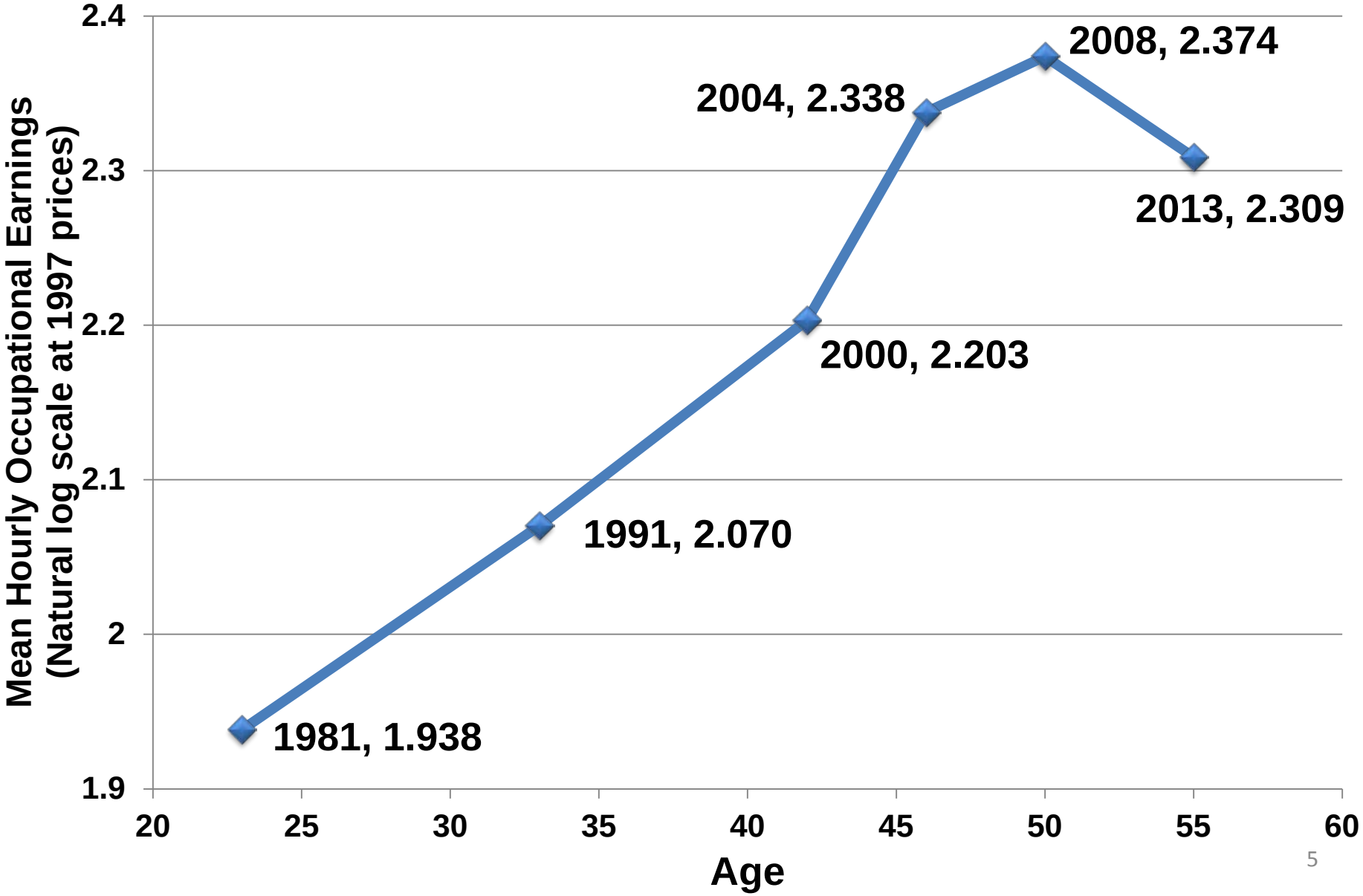
*** PMS and NCDS 1-2. Immigrant-not resident in Great Britain, NCDS4 onwards no address or refusal to participate**

NCDS Survey Response (%) [n = 18558]

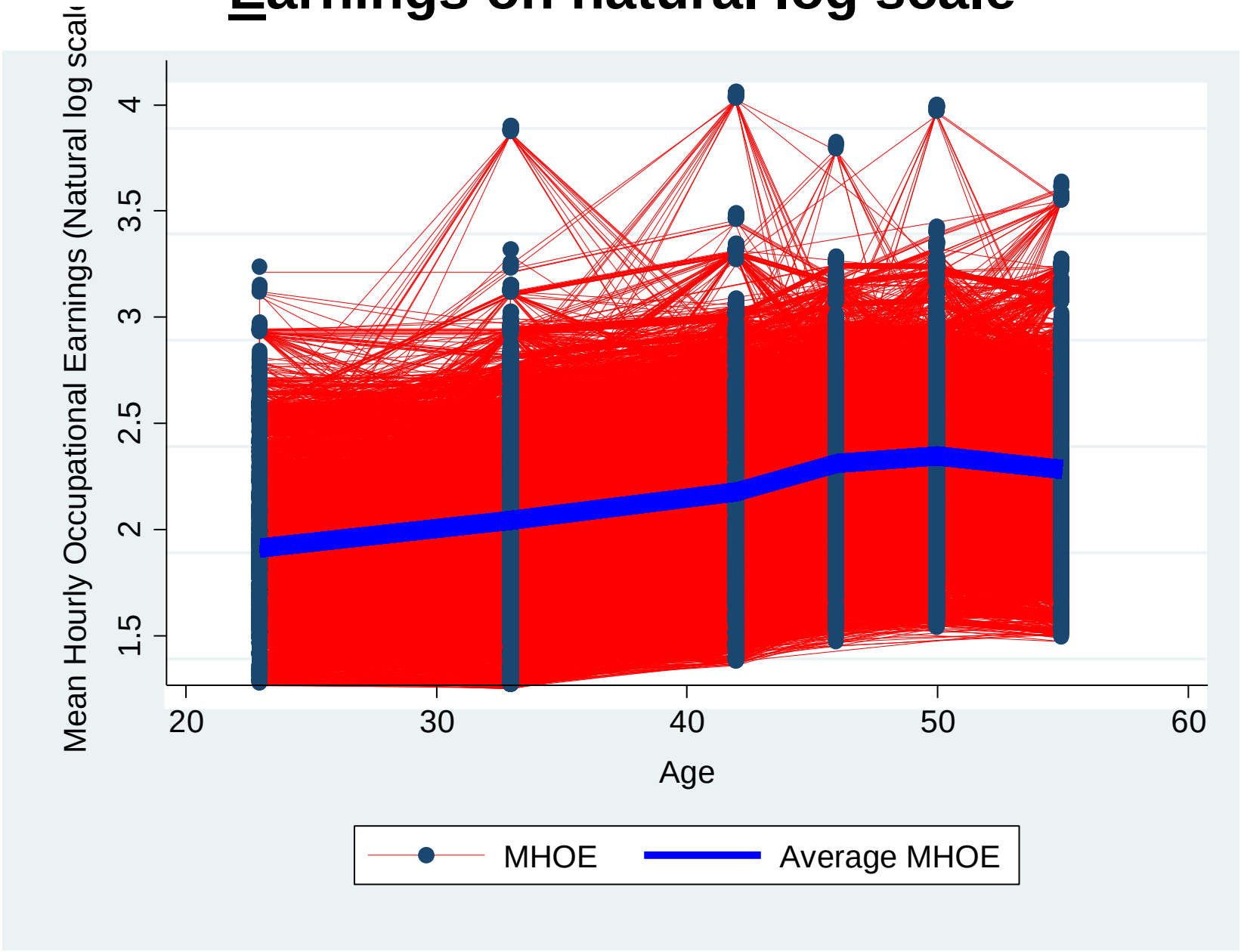
Sweep	PMS	NCDS1	NCDS2	NCDS3	NCDS4	NCDS5	NCDS6	Bio-medical	NCDS7	NCDS8	NCDS9
Year	1958	1965	1969	1974	1981	1991	2000	2002	2004	2008	2013
Age	Birth	7	11	16	23	33	42	44	46	50	55
Productive	93.8	83.1	82.6	79.0	67.6	61.8	61.5	50.5	51.4	52.8	49.2
Refusal	0.0	0.4	4.3	6.2	4.9	7.4	6.2	15.2	7.8	6.5	3.1
Non-contact	1.2	5.6	2.2	4.2	9.0	7.5	9.9	4.3	3.3	4.5	4.6
Other unproductive	0.0	0.9	1.1	1.6	2.2	5.1	1.4	0.2	0.6	1.8	2.6
Ineligible	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.4	0.1	0.4	0.0
Not issued *	5.0	3.0	1.5	0.0	4.6	5.4	7.6	15.7	22.9	19.1	24.5
Emigrant	0.0	2.6	3.8	4.3	6.4	7.2	6.8	6.6	6.9	7.0	6.9
Dead	0.0	4.4	4.5	4.7	5.2	5.7	6.5	7.1	7.1	7.9	8.9
Total	100	100	100	100	100	100	100	100	100	100	100

* PMS and NCDS 1-2. Immigrant-not resident in Great Britain, NCDS4 onwards no address or refusal to participate

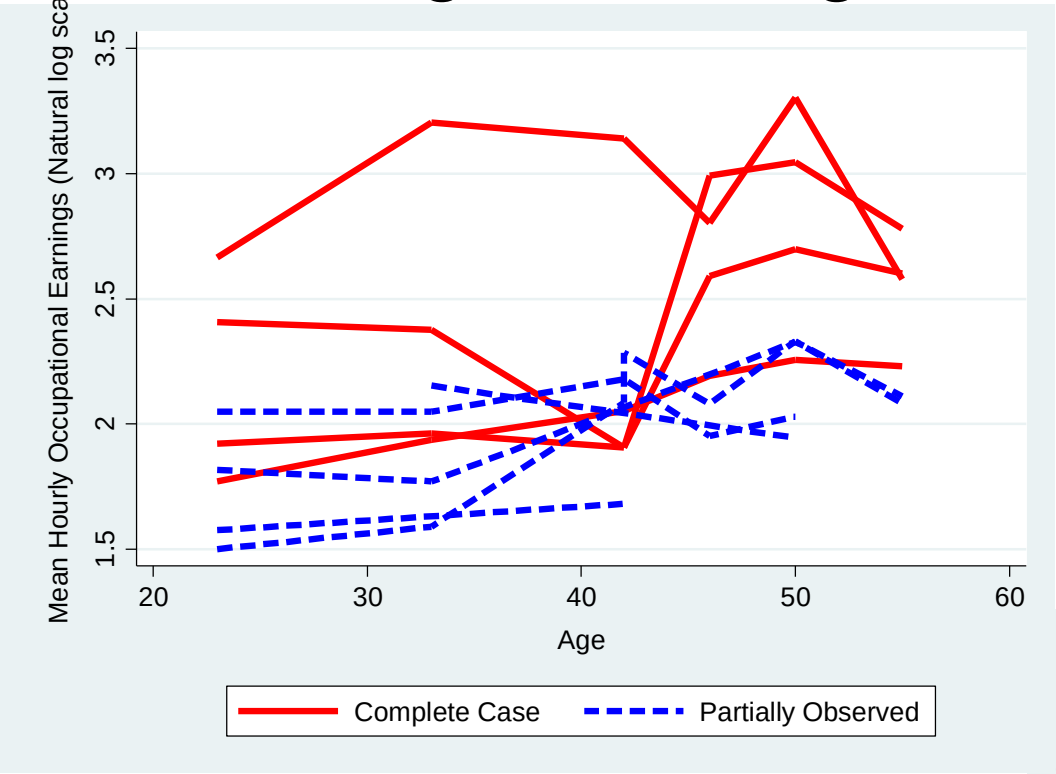
Average Life course Mean Hourly Occupational Earnings on natural log scale



Life course Mean Hourly Occupational Earnings on natural log scale



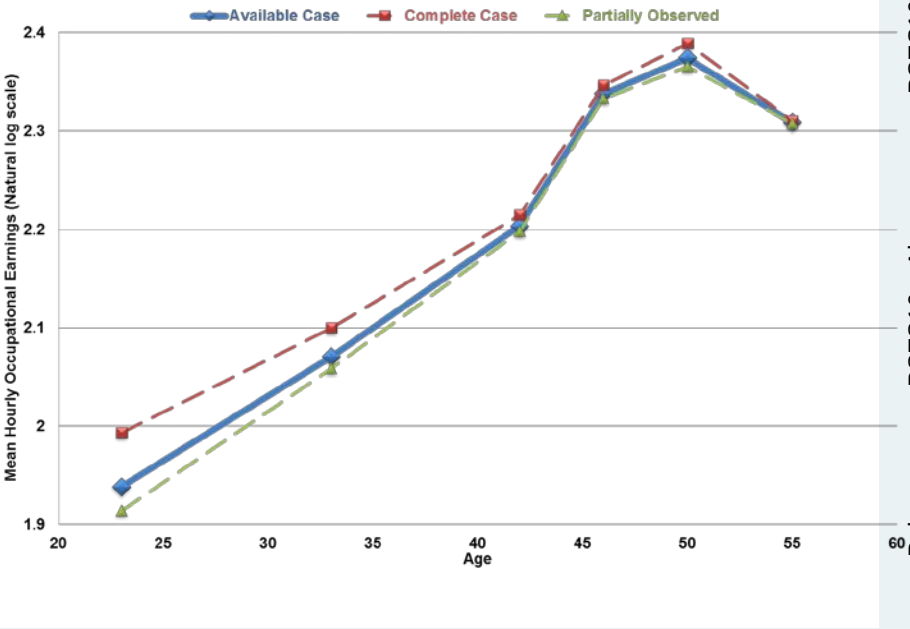
Meaningful missingness in life course MHOE

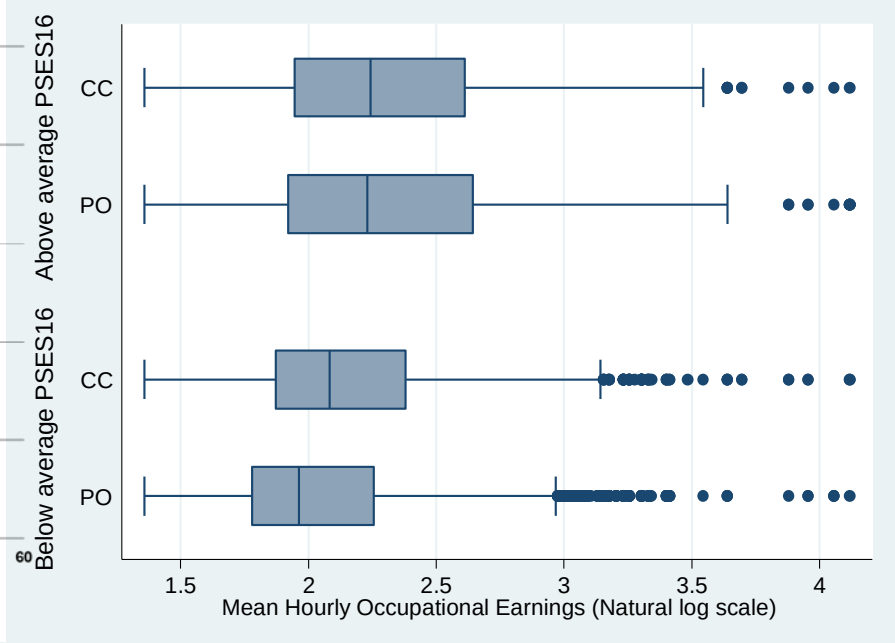


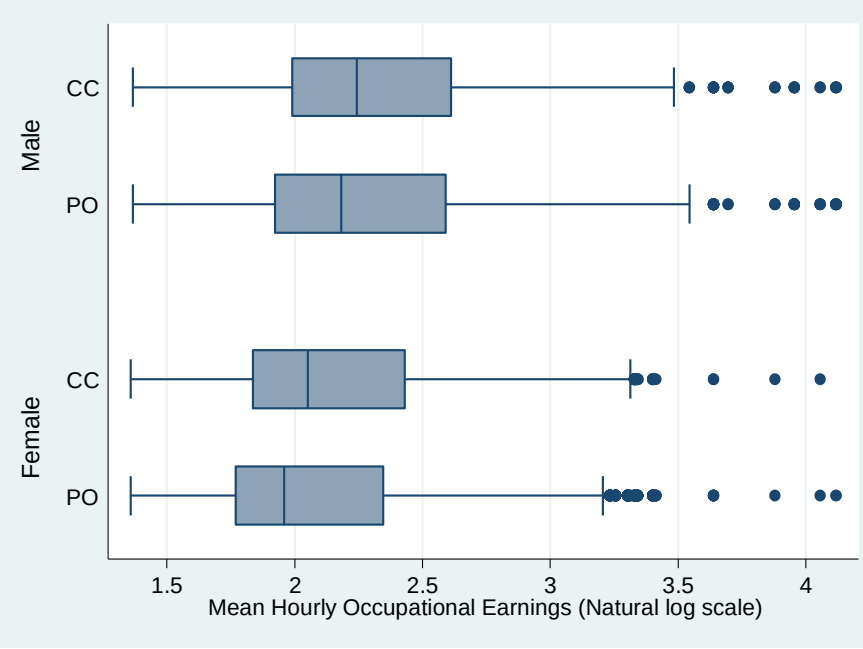
A demonstration using 10 NCDS survey members' life course mean hourly occupational earnings data

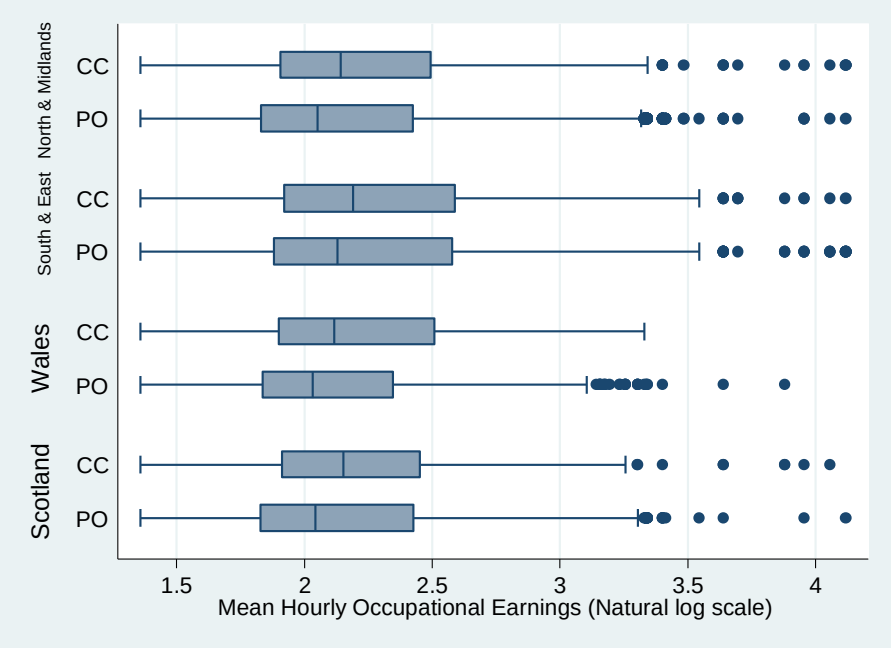
Given the data, the evidence suggests that these data are not Missing Completely at Random (MCAR)

Evidence of Missing at Random (MAR)









Life course step function MLM of interest

$$
\begin{aligned}
&\ln \text{Mean Hourly Occupational Earnings}_{tj} \\
&= \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t \\
&+ \beta_6 \text{age55}_t + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j + \beta_9 \text{age33}_t * \text{PSES16}_j \\
&+ \beta_{10} \text{age42}_t * \text{PSES16}_j + \beta_{11} \text{age46}_t * \text{PSES16}_j \\
&+ \beta_{12} \text{age50}_t * \text{PSES16}_j + \beta_{13} \text{age55}_t * \text{PSES16}_j \\
&+ \beta_{14} \text{age33}_t * \text{Female}_j + \beta_{15} \text{age42}_t * \text{Female}_j \\
&+ \beta_{16} \text{age46}_t * \text{Female}_j + \beta_{17} \text{age50}_t * \text{Female}_j \\
&+ \beta_{18} \text{age55}_t * \text{Female}_j + \beta_{19} \text{South \& East}_{tj} + \beta_{20} \text{Wales}_{tj} \\
&+ \beta_{21} \text{Scotland}_{tj} + u_{1j} \text{age23}_t + u_{2j} \text{age33}_t + u_{3j} \text{age42}_t \\
&+ u_{4j} \text{age46}_t + u_{5j} \text{age50}_t + u_{6j} \text{age55}_t
\end{aligned}
$$

Life course step function MLM of interest

$$
\begin{aligned}
&\ln \text{Mean Hourly Occupational Earnings}_{tj} \\
&= \boldsymbol{\beta_1 age23_t + \beta_2 age33_t + \beta_3 age42_t + \beta_4 age46_t} \\
&+ \boldsymbol{\beta_5 age50_t + \beta_6 age55_t} + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j \\
&+ \beta_9 \text{age33}_t * \text{PSES16}_j + \beta_{10} \text{age42}_t * \text{PSES16}_j \\
&+ \beta_{11} \text{age46}_t * \text{PSES16}_j + \beta_{12} \text{age50}_t * \text{PSES16}_j \\
&+ \beta_{13} \text{age55}_t * \text{PSES16}_j + \beta_{14} \text{age33}_t * \text{Female}_j \\
&+ \beta_{15} \text{age42}_t * \text{Female}_j + \beta_{16} \text{age46}_t * \text{Female}_j \\
&+ \beta_{17} \text{age50}_t * \text{Female}_j + \beta_{18} \text{age55}_t * \text{Female}_j \\
&+ \beta_{19} \text{South \& East}_{tj} + \beta_{20} \text{Wales}_{tj} + \beta_{21} \text{Scotland}_{tj} \\
&+ u_{1j} \text{age23}_t + u_{2j} \text{age33}_t + u_{3j} \text{age42}_t + u_{4j} \text{age46}_t \\
&+ u_{5j} \text{age50}_t + u_{6j} \text{age55}_t
\end{aligned}
$$

Life course step function MLM of interest

$$
\begin{aligned}
&\ln \text{Mean Hourly Occupational Earnings}_{tj} \\
&= \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t \\
&+ \beta_6 \text{age55}_t + \boldsymbol{\beta_7 \text{PSES16}_j} + \boldsymbol{\beta_8 \text{Female}_j} \\
&+ \beta_9 \text{age33}_t * \text{PSES16}_j + \beta_{10} \text{age42}_t * \text{PSES16}_j \\
&+ \beta_{11} \text{age46}_t * \text{PSES16}_j + \beta_{12} \text{age50}_t * \text{PSES16}_j \\
&+ \beta_{13} \text{age55}_t * \text{PSES16}_j + \beta_{14} \text{age33}_t * \text{Female}_j \\
&+ \beta_{15} \text{age42}_t * \text{Female}_j + \beta_{16} \text{age46}_t * \text{Female}_j \\
&+ \beta_{17} \text{age50}_t * \text{Female}_j + \beta_{18} \text{age55}_t * \text{Female}_j \\
&+ \boldsymbol{\beta_{19} \text{South \& East}_{tj}} + \boldsymbol{\beta_{20} \text{Wales}_{tj}} + \boldsymbol{\beta_{21} \text{Scotland}_{tj}} \\
&+ u_{1j} \text{age23}_t + u_{2j} \text{age33}_t + u_{3j} \text{age42}_t + u_{4j} \text{age46}_t \\
&+ u_{5j} \text{age50}_t + u_{6j} \text{age55}_t
\end{aligned}
$$

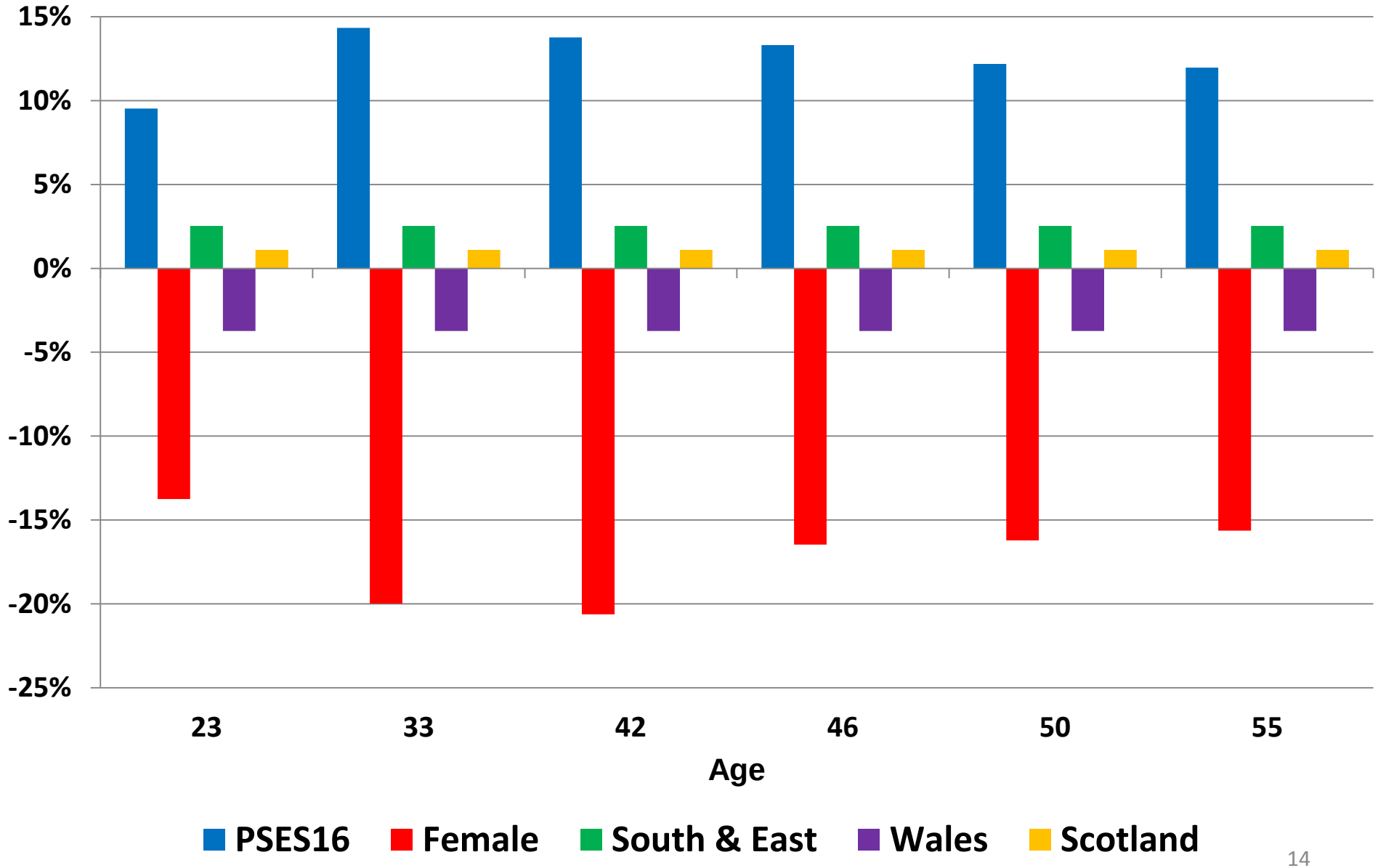
Life course step function MLM of interest

$$
\begin{aligned}
&\ln \text{Mean Hourly Occupational Earnings}_{tj} \\
&= \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t \\
&+ \beta_6 \text{age55}_t + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j \\
&+ \boldsymbol{\beta_9 \text{age33}_t * \text{PSES16}_j + \beta_{10} \text{age42}_t * \text{PSES16}_j} \\
&+ \boldsymbol{\beta_{11} \text{age46}_t * \text{PSES16}_j + \beta_{12} \text{age50}_t * \text{PSES16}_j} \\
&+ \boldsymbol{\beta_{13} \text{age55}_t * \text{PSES16}_j + \beta_{14} \text{age33}_t * \text{Female}_j} \\
&+ \boldsymbol{\beta_{15} \text{age42}_t * \text{Female}_j + \beta_{16} \text{age46}_t * \text{Female}_j} \\
&+ \boldsymbol{\beta_{17} \text{age50}_t * \text{Female}_j + \beta_{18} \text{age55}_t * \text{Female}_j} \\
&+ \beta_{19} \text{South \& East}_{tj} + \beta_{20} \text{Wales}_{tj} + \beta_{21} \text{Scotland}_{tj} \\
&+ u_{1j} \text{age23}_t + u_{2j} \text{age33}_t + u_{3j} \text{age42}_t + u_{4j} \text{age46}_t \\
&+ u_{5j} \text{age50}_t + u_{6j} \text{age55}_t
\end{aligned}
$$

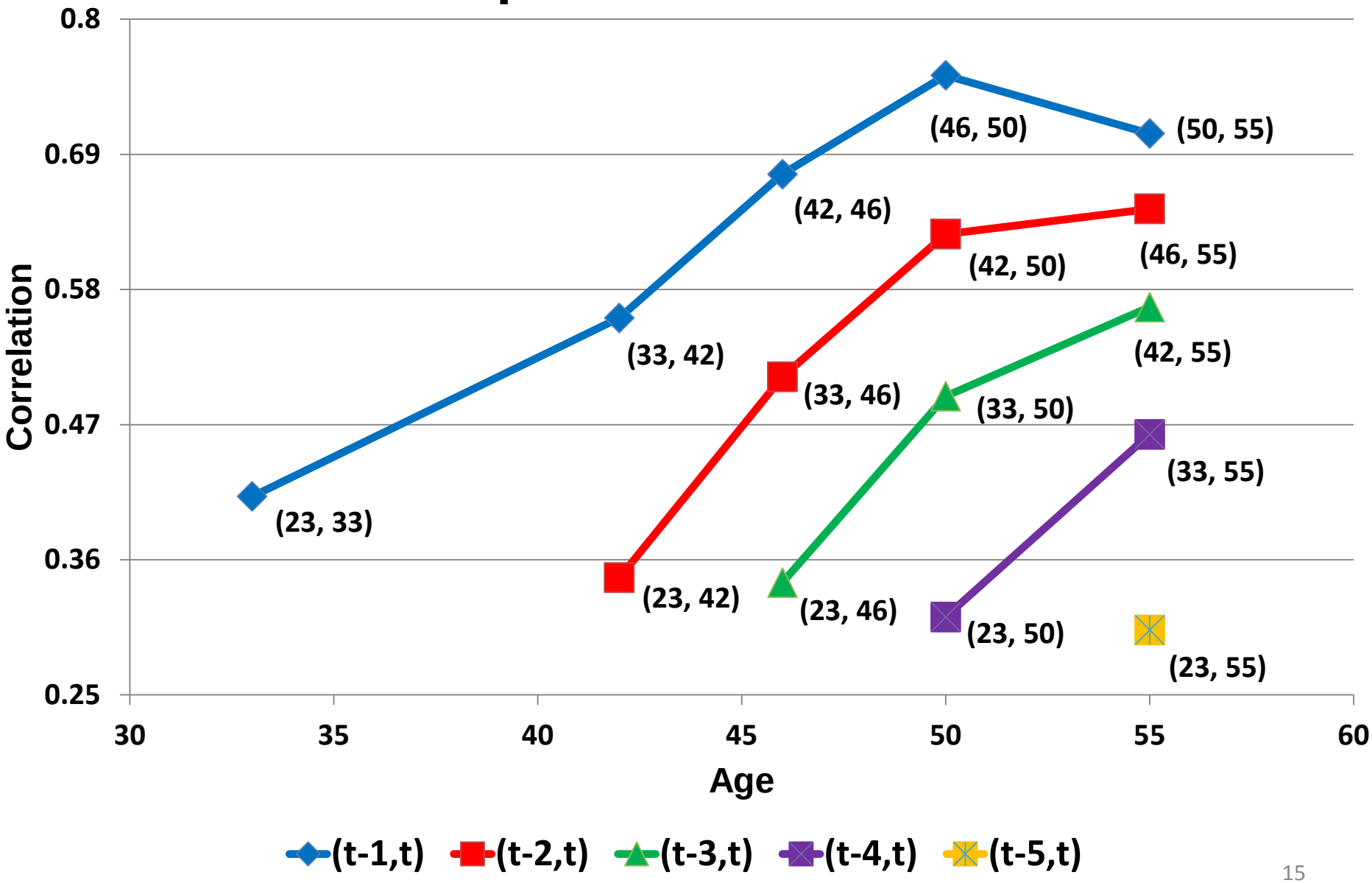
Life course step function MLM of interest

$$\ln \text{Mean Hourly Occupational Earnings}_{tj}$$
$$= \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t$$
$$+ \beta_6 \text{age55}_t + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j + \beta_9 \text{age33}_t * \text{PSES16}_j$$
$$+ \beta_{10} \text{age42}_t * \text{PSES16}_j + \beta_{11} \text{age46}_t * \text{PSES16}_j$$
$$+ \beta_{12} \text{age50}_t * \text{PSES16}_j + \beta_{13} \text{age55}_t * \text{PSES16}_j$$
$$+ \beta_{14} \text{age33}_t * \text{Female}_j + \beta_{15} \text{age42}_t * \text{Female}_j$$
$$+ \beta_{16} \text{age46}_t * \text{Female}_j + \beta_{17} \text{age50}_t * \text{Female}_j$$
$$+ \beta_{18} \text{age55}_t * \text{Female}_j + \beta_{19} \text{South \& East}_{tj} + \beta_{20} \text{Wales}_{tj}$$
$$+ \beta_{21} \text{Scotland}_{tj} + \mathbf{u_{1j} age23_t} + \mathbf{u_{2j} age33_t} + \mathbf{u_{3j} age42_t}$$
$$+ \mathbf{u_{4j} age46_t} + \mathbf{u_{5j} age50_t} + \mathbf{u_{6j} age55_t}$$

Life course step function MLM fixed effects



Life course step function MLM random effect positive correlations



Concluding remarks:

1. NCDS is a fantastic resource but beware of non-response

2. MHOE provides an unbounded measure of SES allowing for more granular investigation of relative inequity

3. Data are not MCAR; there is evidence of MAR but MNAR sensitivity analysis would make investigation more robust

4. Treating time as a step function allows for a non-naïve yet simple interpretation of life course development

Thanks for listening

Any questions?

Email: adrian.byrne@manchester.ac.uk