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

<table>
<thead>
<tr>
<th>Sweep</th>
<th>PMS</th>
<th>NCDS1</th>
<th>NCDS2</th>
<th>NCDS3</th>
<th>NCDS4</th>
<th>NCDS5</th>
<th>NCDS6</th>
<th>Bio-medical</th>
<th>NCDS7</th>
<th>NCDS8</th>
<th>NCDS9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Birth</td>
<td>7</td>
<td>11</td>
<td>16</td>
<td>23</td>
<td>33</td>
<td>42</td>
<td>44</td>
<td>46</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Productive</td>
<td>17415</td>
<td>15425</td>
<td>15337</td>
<td>14654</td>
<td>12537</td>
<td>11469</td>
<td>11419</td>
<td>9377</td>
<td>9534</td>
<td>9790</td>
</tr>
<tr>
<td></td>
<td>Refusal</td>
<td>0</td>
<td>80</td>
<td>797</td>
<td>1151</td>
<td>915</td>
<td>1365</td>
<td>1148</td>
<td>2829</td>
<td>1448</td>
<td>1214</td>
</tr>
<tr>
<td></td>
<td>Non-contact</td>
<td>218</td>
<td>1036</td>
<td>406</td>
<td>786</td>
<td>1675</td>
<td>1394</td>
<td>1832</td>
<td>792</td>
<td>612</td>
<td>835</td>
</tr>
<tr>
<td></td>
<td>Other unproductive</td>
<td>0</td>
<td>173</td>
<td>202</td>
<td>295</td>
<td>413</td>
<td>953</td>
<td>263</td>
<td>31</td>
<td>109</td>
<td>332</td>
</tr>
<tr>
<td></td>
<td>Ineligible</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>65</td>
<td>11</td>
<td>81</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Not issued *</td>
<td>925</td>
<td>548</td>
<td>275</td>
<td>0</td>
<td>862</td>
<td>993</td>
<td>1415</td>
<td>2908</td>
<td>4248</td>
<td>3553</td>
</tr>
<tr>
<td></td>
<td>Emigrant</td>
<td>0</td>
<td>475</td>
<td>701</td>
<td>799</td>
<td>1196</td>
<td>1335</td>
<td>1268</td>
<td>1234</td>
<td>1272</td>
<td>1293</td>
</tr>
<tr>
<td></td>
<td>Dead</td>
<td>0</td>
<td>821</td>
<td>840</td>
<td>873</td>
<td>960</td>
<td>1049</td>
<td>1200</td>
<td>1322</td>
<td>1324</td>
<td>1460</td>
</tr>
<tr>
<td>Total</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
<td>18558</td>
</tr>
</tbody>
</table>

* PMS and NCDS 1-2. Immigrant-not resident in Great Britain, NCDS4 onwards no address or refusal to participate.
## NCDS Survey Response (%) [n = 18558]

<table>
<thead>
<tr>
<th>Sweep</th>
<th>PMS</th>
<th>NCDS1</th>
<th>NCDS2</th>
<th>NCDS3</th>
<th>NCDS4</th>
<th>NCDS5</th>
<th>NCDS6</th>
<th>Biomedical</th>
<th>NCDS7</th>
<th>NCDS8</th>
<th>NCDS9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1958</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1969</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth</td>
<td>7</td>
<td>11</td>
<td>16</td>
<td>23</td>
<td>33</td>
<td>42</td>
<td>44</td>
<td>46</td>
<td>50</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>19 Productive</td>
<td>93.8</td>
<td>83.1</td>
<td>82.6</td>
<td>79.0</td>
<td>67.6</td>
<td>61.8</td>
<td>61.5</td>
<td>50.5</td>
<td>51.4</td>
<td>52.8</td>
<td>49.2</td>
</tr>
<tr>
<td>Refusal</td>
<td>0.0</td>
<td>0.4</td>
<td>4.3</td>
<td>6.2</td>
<td>4.9</td>
<td>7.4</td>
<td>6.2</td>
<td>15.2</td>
<td>7.8</td>
<td>6.5</td>
<td>3.1</td>
</tr>
<tr>
<td>Non-contact</td>
<td>1.2</td>
<td>5.6</td>
<td>2.2</td>
<td>4.2</td>
<td>9.0</td>
<td>7.5</td>
<td>9.9</td>
<td>4.3</td>
<td>3.3</td>
<td>4.5</td>
<td>4.6</td>
</tr>
<tr>
<td>Other unproductive</td>
<td>0.0</td>
<td>0.9</td>
<td>1.1</td>
<td>1.6</td>
<td>2.2</td>
<td>5.1</td>
<td>1.4</td>
<td>0.2</td>
<td>0.6</td>
<td>1.8</td>
<td>2.6</td>
</tr>
<tr>
<td>Ineligible</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.4</td>
<td>0.1</td>
<td>0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Not issued *</td>
<td>5.0</td>
<td>3.0</td>
<td>1.5</td>
<td>0.0</td>
<td>4.6</td>
<td>5.4</td>
<td>7.6</td>
<td>15.7</td>
<td>22.9</td>
<td>19.1</td>
<td>24.5</td>
</tr>
<tr>
<td>Emigrant</td>
<td>0.0</td>
<td>2.6</td>
<td>3.8</td>
<td>4.3</td>
<td>6.4</td>
<td>7.2</td>
<td>6.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.0</td>
<td>6.9</td>
</tr>
<tr>
<td>Dead</td>
<td>0.0</td>
<td>4.4</td>
<td>4.5</td>
<td>4.7</td>
<td>5.2</td>
<td>5.7</td>
<td>6.5</td>
<td>7.1</td>
<td>7.1</td>
<td>7.9</td>
<td>8.9</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

* 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

Mean Hourly Occupational Earnings (Natural log scale at 1997 prices)

Age

1981, 1.938
1991, 2.070
2000, 2.203
2004, 2.338
2008, 2.374
2013, 2.309
Life course Mean Hourly Occupational Earnings on natural log scale

Mean Hourly Occupational Earnings (Natural log scale)

Age

MHOE Average MHOE
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)

- Available Case
- Complete Case
- Partially Observed

Mean Hourly Occupational Earnings (Natural log scale)

- Below average PSES16
- Above average PSES16

PO
CC

Female
Male

Scotland
North & Midlands
South & East

1.5 2 2.5 3 3.5 4

1.5 2 2.5 3 3.5 4

1.5 2 2.5 3 3.5 4
Life course step function MLM of interest

\[
\ln \text{Mean Hourly Occupational Earnings}_{tj} = \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t \\
+ \beta_6 \text{age55}_t + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j + \beta_9 \text{age33}_t \ast \text{PSES16}_j \\
+ \beta_{10} \text{age42}_t \ast \text{PSES16}_j + \beta_{11} \text{age46}_t \ast \text{PSES16}_j \\
+ \beta_{12} \text{age50}_t \ast \text{PSES16}_j + \beta_{13} \text{age55}_t \ast \text{PSES16}_j \\
+ \beta_{14} \text{age33}_t \ast \text{Female}_j + \beta_{15} \text{age42}_t \ast \text{Female}_j \\
+ \beta_{16} \text{age46}_t \ast \text{Female}_j + \beta_{17} \text{age50}_t \ast \text{Female}_j \\
+ \beta_{18} \text{age55}_t \ast \text{Female}_j + \beta_{19} \text{South & East}_t + \beta_{20} \text{Wales}_t \\
+ \beta_{21} \text{Scotland}_t + u_1 j \text{age23}_t + u_2 j \text{age33}_t + u_3 j \text{age42}_t \\
+ u_4 j \text{age46}_t + u_5 j \text{age50}_t + u_6 j \text{age55}_t
\]
Life course step function MLM of interest

\[
\ln \text{Mean Hourly Occupational Earnings}_{tj} = \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t + \beta_6 \text{age55}_t + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j + \beta_9 \text{age33}_t \times \text{PSES16}_j + \beta_{10} \text{age42}_t \times \text{PSES16}_j + \beta_{11} \text{age46}_t \times \text{PSES16}_j + \beta_{12} \text{age50}_t \times \text{PSES16}_j + \beta_{13} \text{age55}_t \times \text{PSES16}_j + \beta_{14} \text{age33}_t \times \text{Female}_j + \beta_{15} \text{age42}_t \times \text{Female}_j + \beta_{16} \text{age46}_t \times \text{Female}_j + \beta_{17} \text{age50}_t \times \text{Female}_j + \beta_{18} \text{age55}_t \times \text{Female}_j + \beta_{19} \text{South & East}_t + \beta_{20} \text{Wales}_t + \beta_{21} \text{Scotland}_t + u_1j \text{age23}_t + u_2j \text{age33}_t + u_3j \text{age42}_t + u_4j \text{age46}_t + u_5j \text{age50}_t + u_6j \text{age55}_t
\]
Life course step function MLM of interest

\[ \ln \text{Mean Hourly Occupational Earnings}_{tj} = \beta_1 \text{age23}_t + \beta_2 \text{age33}_t + \beta_3 \text{age42}_t + \beta_4 \text{age46}_t + \beta_5 \text{age50}_t + \beta_6 \text{age55}_t + \beta_7 \text{PSES16}_j + \beta_8 \text{Female}_j + \beta_9 \text{age33}_t \ast \text{PSES16}_j + \beta_{10} \text{age42}_t \ast \text{PSES16}_j + \beta_{11} \text{age46}_t \ast \text{PSES16}_j + \beta_{12} \text{age50}_t \ast \text{PSES16}_j + \beta_{13} \text{age55}_t \ast \text{PSES16}_j + \beta_{14} \text{age33}_t \ast \text{Female}_j + \beta_{15} \text{age42}_t \ast \text{Female}_j + \beta_{16} \text{age46}_t \ast \text{Female}_j + \beta_{17} \text{age50}_t \ast \text{Female}_j + \beta_{18} \text{age55}_t \ast \text{Female}_j + \beta_{19} \text{South \& East}_{tj} + \beta_{20} \text{Wales}_{tj} + \beta_{21} \text{Scotland}_{tj} + u_{1j} \text{age23}_t + u_{2j} \text{age33}_t + u_{3j} \text{age42}_t + u_{4j} \text{age46}_t + u_{5j} \text{age50}_t + u_{6j} \text{age55}_t \]
Life course step function MLM of interest

\[
\ln \text{Mean Hourly Occupational Earnings}_{tj} = \beta_{1}\text{age23}_t + \beta_{2}\text{age33}_t + \beta_{3}\text{age42}_t + \beta_{4}\text{age46}_t + \beta_{5}\text{age50}_t \\
+ \beta_{6}\text{age55}_t + \beta_{7}\text{PSES16}_j + \beta_{8}\text{Female}_j \\
+ \beta_{9}\text{age33}_t \times \text{PSES16}_j + \beta_{10}\text{age42}_t \times \text{PSES16}_j \\
+ \beta_{11}\text{age46}_t \times \text{PSES16}_j + \beta_{12}\text{age50}_t \times \text{PSES16}_j \\
+ \beta_{13}\text{age55}_t \times \text{PSES16}_j + \beta_{14}\text{age33}_t \times \text{Female}_j \\
+ \beta_{15}\text{age42}_t \times \text{Female}_j + \beta_{16}\text{age46}_t \times \text{Female}_j \\
+ \beta_{17}\text{age50}_t \times \text{Female}_j + \beta_{18}\text{age55}_t \times \text{Female}_j \\
+ \beta_{19}\text{South & East}_t + \beta_{20}\text{Wales}_t + \beta_{21}\text{Scotland}_t \\
+ u_{1j}\text{age23}_t + u_{2j}\text{age33}_t + u_{3j}\text{age42}_t + u_{4j}\text{age46}_t \\
+ u_{5j}\text{age50}_t + u_{6j}\text{age55}_t
\]
Life course step function MLM of interest

\[ \ln \text{Mean Hourly Occupational Earnings}_{tj} = \beta_1 \text{age}^{23}_t + \beta_2 \text{age}^{33}_t + \beta_3 \text{age}^{42}_t + \beta_4 \text{age}^{46}_t + \beta_5 \text{age}^{50}_t \\
+ \beta_6 \text{age}^{55}_t + \beta_7 \text{PSES}_{16j} + \beta_8 \text{Female}_j + \beta_9 \text{age}^{33}_t \ast \text{PSES}_{16j} \\
+ \beta_{10} \text{age}^{42}_t \ast \text{PSES}_{16j} + \beta_{11} \text{age}^{46}_t \ast \text{PSES}_{16j} \\
+ \beta_{12} \text{age}^{50}_t \ast \text{PSES}_{16j} + \beta_{13} \text{age}^{55}_t \ast \text{PSES}_{16j} \\
+ \beta_{14} \text{age}^{33}_t \ast \text{Female}_j + \beta_{15} \text{age}^{42}_t \ast \text{Female}_j \\
+ \beta_{16} \text{age}^{46}_t \ast \text{Female}_j + \beta_{17} \text{age}^{50}_t \ast \text{Female}_j \\
+ \beta_{18} \text{age}^{55}_t \ast \text{Female}_j + \beta_{19} \text{South} \& \text{East}_{tj} + \beta_{20} \text{Wales}_{tj} \\
+ \beta_{21} \text{Scotland}_{tj} + u_{1j} \text{age}^{23}_t + u_{2j} \text{age}^{33}_t + u_{3j} \text{age}^{42}_t \\
+ u_{4j} \text{age}^{46}_t + u_{5j} \text{age}^{50}_t + u_{6j} \text{age}^{55}_t \]
Life course step function MLM fixed effects
Life course step function MLM random effect positive correlations

Correlation

Age

0.25 0.36 0.47 0.58 0.69 0.8
30 35 40 45 50 55 60

(t-1,t) (t-2,t) (t-3,t) (t-4,t) (t-5,t)
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