Estimates of the Causal Effect of Education on Earnings over the Lifecycle

UK evidence of a Non-separable Specification with Cohort Effects and Endogenous Education

Giuseppe Migali
and
Ian Walker
Lancaster University Management School
Motivation

• Huge industry estimating workhorse HCEF
  – \( \log w_i = \mathbf{X}_i d + a.\text{Age}_i + b.\text{Age}_i^2 + c.S_i + u_i \)

• Substantial work on case where \( \text{cov}(S,u) \neq 0 \)
  – But special case of \( g(w, \text{Age}, S, \mathbf{X}) + u = 0 \)

• Strong separability restriction on \( g \)
  – Lifecycle log \( w \) profiles parallel in \( S \)
    • OLS research rejects this, although no IV estimates

• Likely to be strong cohort effects in wages
  – Cohort effects may be systematic

• Hard to separate cohort from lifecycle effects
  – Failure to control likely to bias lifecycle estimates
Existing literature

• Heckman, Lochner and Todd (JHC 2008)
  – Estimates $g(w, Age, S, X)$ nonparametrically
  – Datasets are large so many $Age*S$ cells
    • But ignores cohort effects
  – Ignores potential endogeneity of $S$
    • Uses crude matching

• But, difficult to use a completely flexible specification in $S$, if $S$ is endogenous
  – Large datasets, but usually few instruments
UK education data

• UK has a straightforward way to group $S$ data by qualifications (5, 4, 3, 2, 1, 0):
  – PG, UG, A-level, GCSE, Sub GCSE, and (Nil)
  – together with vocational equivalents

• So collapse $S$ into small groups
  – Estimate HCEF by $S$ group
  – Tractable, but no compromise on generality
    • Retain nonparametric effects of (grouped) $S$
      – But ordering in $S$ groups helps economise on instruments
    • Retain nonparametric in $Age$ if desired
      – or group $Age$ into bands
      – or parameterise as a continuous function
UK education data

• NCDS data on qualifications
  – Multiple treatments
  – But ordered
• Matching helps alot (Blundell et al, JRSS 2005)
• But needs rich data
  – Family background and test scores
• In the absence of rich data
  – Blundell et al suggest exploiting available exclusion restrictions to estimate control function
• (o)Heckman selection correction
Endogenous S

• Each S group needs a separate IV
• And IV provides only LATE
  – Not generally comparable with OLS
• Alternatively - order S and adopt Heckit
  – then we can compare with OLS
  – and economise on exclusion restrictions
  – at the cost of joint normality (alone)
• RoSLA and Month of Birth exclusions
  – Born after August 1958 (Harmon/Walker 1995)
  – Sept = 12, .... Aug = 1 (Crawford et al 2007)
Cohort effects

• Strong cohort effects in $S$ in UK
  – And probably in wages too

• But age and year of birth are perfectly collinear in a cross section
  – Highly collinear even in pooled LFS x-sections

• LFS is (since 97) a short 5-quarter panel
  – Use the panel to identify the lifecycle effect
    • Estimate $\Delta \log w = a + 2b \cdot Age$ by $S$ group
  – And use the cross section variation in year of birth to identify cohort effects
    • Estimate other HCEF coeffs by $S$
Method

• Not possible to pool LLFS and QLFS
  – no common id
  – So cannot use SURE with x-equation restrictions
• QLFS too short to separate cohort/lifecycle
• Use LLFS 97-08 to estimate $\Delta w$ equations
  – $\Delta \log w_{is} = a_{is} + 2b.Age_{is} + u_{is}$ for $s=1..5$
  – Unbiased $a_s$ and $b_s$ if “ability” is a FE and additive
  – Cohort effects drop out if separable
• Impose these estimates on levels equations
  – Estimate these using QLFS by oheckman - by S
  – Allow for additive cohort effects
RoSLA

- Chevalier et al. *EJ* 2004 shows RoSLA affects only bottom of S distribution
Month of Birth

- Youngest in class do worse (Crawford et al IFS WP 2008)
  - Entry, peer, and developmental effects
Raw QLFS pooled data:
OLS quadratics with discrete S groups and no cohort effects

Male

Female
First stage results
Marginal effects from Ordered Probit

- Collapse NVQs into 0/1, 2, 3, 4/5
- Includes cubic in year of birth
  - So RoSLA is a RD

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th></th>
<th>Females</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>month of birth</td>
<td>-0.00164*</td>
<td>-0.00371***</td>
<td>(0.00095)</td>
<td>(0.00090)</td>
</tr>
<tr>
<td></td>
<td>(0.02464)</td>
<td>(0.02247)</td>
<td>(0.01523)</td>
<td>(0.01424)</td>
</tr>
<tr>
<td>non white</td>
<td>-0.15139***</td>
<td>-0.25877***</td>
<td>(0.05980***</td>
<td>(0.07458***</td>
</tr>
<tr>
<td></td>
<td>(0.01523)</td>
<td>(0.01424)</td>
<td>(0.03444)</td>
<td>(0.03353)</td>
</tr>
<tr>
<td>year of birth</td>
<td>0.39362***</td>
<td>0.48996***</td>
<td>(0.03444)</td>
<td>(0.03353)</td>
</tr>
<tr>
<td></td>
<td>(0.00059)</td>
<td>(0.00057)</td>
<td>(0.00000)</td>
<td>(0.00000)</td>
</tr>
<tr>
<td>( (year of birth)^2 )</td>
<td>-0.00631***</td>
<td>-0.00753***</td>
<td>(0.00059)</td>
<td>(0.00057)</td>
</tr>
<tr>
<td></td>
<td>(0.00000)</td>
<td>(0.00000)</td>
<td>(0.00000)</td>
<td>(0.00000)</td>
</tr>
<tr>
<td>( (year of birth)^3 )</td>
<td>0.00003***</td>
<td>0.00004***</td>
<td>(0.00000)</td>
<td>(0.00000)</td>
</tr>
</tbody>
</table>

Table 5: First Step - ordered probit - mfx

<table>
<thead>
<tr>
<th>Dep var: NVQ levels</th>
<th>Males RoSLA</th>
<th>month of birth</th>
<th>Females RoSLA</th>
<th>month of birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVQ0</td>
<td>0.01059</td>
<td>0.00029</td>
<td>0.01550</td>
<td>0.00078</td>
</tr>
<tr>
<td></td>
<td>(0.00267)</td>
<td>(0.00017)</td>
<td>(0.00293)</td>
<td>(0.00019)</td>
</tr>
<tr>
<td>NVQ1</td>
<td>0.00356</td>
<td>0.00010</td>
<td>0.00545</td>
<td>0.00027</td>
</tr>
<tr>
<td></td>
<td>(0.00090)</td>
<td>(0.00006)</td>
<td>(0.00104)</td>
<td>(0.00007)</td>
</tr>
<tr>
<td>NVQ2</td>
<td>0.00786</td>
<td>0.00021</td>
<td>0.00878</td>
<td>0.00043</td>
</tr>
<tr>
<td></td>
<td>(0.00201)</td>
<td>(0.00012)</td>
<td>(0.00171)</td>
<td>(0.00010)</td>
</tr>
<tr>
<td>NVQ3</td>
<td>0.00088</td>
<td>-0.00000</td>
<td>-0.00301</td>
<td>-0.00016</td>
</tr>
<tr>
<td></td>
<td>(0.00008)</td>
<td>(0.00000)</td>
<td>(0.00005)</td>
<td>(0.00004)</td>
</tr>
<tr>
<td>NVQ4</td>
<td>-0.02207</td>
<td>-0.00060</td>
<td>-0.02672</td>
<td>-0.00132</td>
</tr>
<tr>
<td></td>
<td>(0.00564)</td>
<td>(0.00035)</td>
<td>(0.00513)</td>
<td>(0.00032)</td>
</tr>
</tbody>
</table>

Estimates significant at 1% std. err. in brackets.
Lifecycle effects
Quadratic in age, assuming no cohort differences

Male

Female
Lifecycle effects
Allowing for cohort differences

Male

Female
Conclusion

• New benchmark HCEF
  – Flexible age, cohort and education effects
  – Yet endogenous education

• LFS estimates
  – Exploit panel element of data to distinguish lifecycle from cohort effects
    • Better with longer panel (BHPS)
  – Exploit month of birth and RoSLA as exclusion restrictions to estimate levels equation
Conclusion

• Strong age effects throughout lifecycle for all $S$ if we impose no cohort effects
  – But allowing for cohort effects, we find very flat age earnings profiles for all $S$ types
  – Strong cohort effects in the data

• Strong ATE of HE on earnings
  – But smaller for more recent cohorts
    • Especially for men

• Decomposing HE effects by “major”
  – Walker and Zhu find major “major” differences
    • ELM trumps STEM