The cohort size-wage relationship in Europe

3rd European User Conference for EU-LFS and EU-SILC
Mannheim, 21-22 March 2013

John Moffat (Swansea University)
Duncan Roth (Philipps-Universität Marburg)
Motivation

• Changes in Europe’s demographic and educational composition:
  – Share of young individuals declining
  – Tendency towards more education

• Age-education structures modelled by relative cohort size

• Identification of causal effect of cohort size on wages:
  – Use of instrumental variable estimation
  – Use of large and recent European data set
  – Identification of effect on individual rather than average wages
Research questions

• Does cohort size have a negative effect on individual wages?

• Are any negative effects permanent of temporary?

• Does cohort size affect the wages of education-gender groups differently?

• How do the estimated cohort size effects compare to the results of related research?
Literature review

• Large cohorts start entering US labour market from 1960s onwards:
  – Effects on wages, unemployment, education, etc.

• Results of empirical studies (mainly from the US):
  – Members of larger cohorts receive reduced earnings
  – Depressing effects larger for highly educated individuals
  – Uncertainty regarding persistence of negative effects
  – Evidence for Europe limited

• Theoretical framework:
  – Identically educated individuals of different age only imperfectly substitutable
  – Highly educated individuals less substitutable across age
Theoretical model (based on Brunello 2010)

- **Education-specific production function:**
  \[ Y_{te} = \phi_{te} N_{te} \]

- **Labour input is a weighted average of age-specific inputs:**
  \[ N_{te} = \left[ \sum_{j=1}^{K} (v_{jte} N_{jte}^{\eta^e}) \right]^{1/\eta^e} \]

- **Firm’s profits:**
  \[ \pi(N_{jte}) = p_{te} \phi_{te} \left[ \sum_{j=1}^{K} (v_{jte} N_{jte}^{\eta^e}) \right]^{1/\eta^e} - \sum_{j=1}^{K} w_{jte} N_{jte} \]

- **Inverse age-education specific labour demand functions:**
  \[ \ln w_{jte} = \ln p_{te} + \ln \phi_{te} + \ln v_{jte} + (\eta^e - 1) \ln \left[ \frac{N_{jte}}{N_{te}} \right] \]
Theoretical model

• Key parameter:
  – Parameter $\eta^e$ reflects the substitutability of identically educated workers of different age
  
  – By definition, $1 > \eta^e > -\infty$ ($\eta=1$ represents perfect substitutability)
  
  – Theoretically, coefficient on relative cohort size expected to be negative if workers are not perfectly substitutable
  
  – Differences in coefficient estimates across educational groups represent differences in elasticity of substitution
Data source and data selection

• Data:
  – EU-SILC Longitudinal UDB 2010 (version 1)
  – Up to 24 countries for period 2007-2010 (20 countries finally used)
• Separate regressions by International Standard Classification of Education:
  – ISCED 2: Lower secondary education
  – ISCED 3: (Upper) secondary education
  – ISCED 5: Tertiary education (also includes ISCED 6)
• Separate regressions for males and females
• Restrictions:
  – Age 20-45 (ISCED 2 and 3)
  – Age 25-45 (ISCED 5)
  – Employed and unemployed individuals
• Regional classification:
  – First level of Nomenclature of Territorial Units for Statistics (NUTS 1)
Empirical model

- **Regression model:**
  \[
  \ln w_{ijkte} = \alpha_1 \ln CS_{jkte} + \alpha_2 \ln[CS_{jkte}] Age_j + \alpha_3 \ln[CS_{jkte}] Age_j^2 + \beta X_{ijkte} + \varepsilon_{ijkte}
  \]

- **Marginal effects of cohort size:**
  \[
  \frac{\partial \ln w_{ijkte}}{\partial \ln CS_{jkte}} = \alpha_1 + \alpha_2 Age_j + \alpha_3 Age_j^2
  \]

- **Key regressor:**
  \[
  CS_{jkte} = \frac{(1/9)N_{j-2kte} + (2/9)N_{j-1kte} + (3/9)N_{jkte} + (4/9)N_{j+1kte} + (1/9)N_{j+2kte}}{N_{kte}}
  \]

- **Weighting:**
  - Construction of cohort size variable
  - Estimation
Empirical model – Control variables

• Age and squared age (age-specific)
  – The age variable is rescaled to 0-25 (ISCED 2, 3) and 0-20 (ISCED 5)

• Time fixed effects (time-specific)

• Geographical area fixed effects (area-specific)

• Employed, occupation, married, part-time, health limitation (individual-by-time-specific)

• Interaction of age and time fixed effects (age-by-time-specific)

• Unemployment rate, degree of urbanisation, interaction of area effects and time fixed effects, interaction of unemployment rate and time fixed effects (area-by-time-specific)
Identification

• Endogeneity of cohort size variable:
  – Self-selection into specific educational groups
  – Self-selection into specific geographic areas
  – Cohort size variables potentially correlated with the error term
    ➔ Least Squares estimation produces inconsistent cohort size coefficients
      (i.e. no convergence in probability on the “true” population value)

• Use of instrumental variables (IV) estimation necessary to identify causal effect of cohort size on wages:
  – Birth rate at year of birth of a specific age cohort
  – Interaction of age (squared age) and lagged birth rate
Identification

- Wages are modelled as the sum of the conditional mean and the error term:
  \[ w_{ijkte} = E[w_{ijkte}|CS, X] + e_{ijkte} = E[w_{ijkte}|CS, X] + \eta_i + u_{ijkte} \]
- If individual do self-select themselves into economically attractive areas and educational groups, then cohort size is itself dependant on the wage rate:
  \[ w_{ijkte} = E[w_{ijkte}|CS(w_{ijkte}), X] + \eta_i + u_{ijkte} \]
- This relationship makes cohort size indirectly depend on the control variables as well as the time-invariant and the idiosyncratic error terms:
  \[ w_{ijkte} = E[w_{ijkte}|CS(X, \eta_i, u_{ijkte}), X] + \eta_i + u_{ijkte} \]

Assumption 1: Cohort size and the (idiosyncratic) error term are positively correlated
  - The larger the error term, the larger the actual wage rate
  - The higher the wage rate, the more individuals will be attracted into the corresponding area/educational group and thus the larger will the cohort size variable be
Identification

• (Positive) Correlation between cohort size variable and error term
  – OLS estimation will produce an inconsistent cohort size coefficient
  – Cohort size coefficient will confound the direct effect on wages from a change in cohort size and the indirect effect stemming from the corresponding change in the error term

Assumption 2: OLS estimation overstates the direct effect of cohort size on wages due to confounding
Model selection

- Availability of panel data allows the choice between two different types of models for the data generating process (dgp):
  - Population-averaged (pooled) model: No individual-specific effects
  - Fixed effects (FE) model: Individual-specific effects that may be correlated with regressors
  - Random effects (RE) model: Individual-specific effects that must not be correlated with regressors

- Choice of model for dgp:
  - Estimation based on FE model attractive as it purges endogeneity that is due to correlation with time-invariant error component
  - However, problem of potential correlation with idiosyncratic error component not solved by use of FE model
  - Possibility of correlation between cohort size regressor and idiosyncratic error term components implies that each model has to be estimated through use of an IV estimator
Model selection

• Choice of model for dgp [continued]:
  – The chosen instrument is time-invariant which rules out the possibility of IV-estimation using the FE model
  – The instrument can be assumed to be uncorrelated with the time-invariant and the idiosyncratic components of the error term ➔ the FE model is not required to identify the causal effect of cohort size on wages (even if time-invariant individual-specific effects are present)
  – IV-estimation is performed for the pooled model

• Limitations of the instrument:
  – The instrument does not vary across all of the dimensions of the endogenous regressor (no variation across education)
  – No regional birth rates available ➔ Loss of efficiency at regional level
  – Limited correlation with cohort size if migration in current period is large
Results I – Relevance of cohort size in wage determination

- Vast majority of cohort size-related variables highly significant for males and ISCED 3 females

- Cohort size variables are less significant (individually and jointly) in the determination of female wages at ISCED 2 and ISCED 5 level

<table>
<thead>
<tr>
<th></th>
<th>ISCED 2f lwage</th>
<th>ISCED 2m lwage</th>
<th>ISCED 3f lwage</th>
<th>ISCED 3m lwage</th>
<th>ISCED 5f lwage</th>
<th>ISCED 5m lwage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICS</td>
<td>0.6822**</td>
<td>0.1218</td>
<td>0.6591***</td>
<td>0.2279</td>
<td>0.8587</td>
<td>1.4368***</td>
</tr>
<tr>
<td></td>
<td>(0.2662)</td>
<td>(0.1538)</td>
<td>(0.2288)</td>
<td>(0.2168)</td>
<td>(0.5894)</td>
<td>(0.4688)</td>
</tr>
<tr>
<td>ICS*Age_20</td>
<td>-0.0797*</td>
<td>-0.0778***</td>
<td>-0.0914***</td>
<td>-0.1241***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0446)</td>
<td>(0.0259)</td>
<td>(0.0242)</td>
<td>(0.0203)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICS*Age_20_sq</td>
<td>0.0022</td>
<td>0.0032***</td>
<td>0.0035***</td>
<td>0.0064***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0018)</td>
<td>(0.0011)</td>
<td>(0.0011)</td>
<td>(0.0009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICS*Age_25</td>
<td>-0.6649*</td>
<td></td>
<td>-0.7344***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.4016)</td>
<td></td>
<td>(0.2555)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICS*Age_25_sq</td>
<td>0.0385*</td>
<td>0.0394***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0232)</td>
<td>(0.0141)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>8,420</td>
<td>13,080</td>
<td>36,638</td>
<td>48,624</td>
<td>31,066</td>
<td>24,346</td>
</tr>
<tr>
<td>R2</td>
<td>0.5984</td>
<td>0.6616</td>
<td>0.7492</td>
<td>0.7838</td>
<td>0.6185</td>
<td>0.6205</td>
</tr>
</tbody>
</table>
Results II – Marginal effects

• Cohort size reduces male wages

• Results less conclusive for female wages

• Future research:
  – Relevance of cohort size for female labour market participation
Results II – Marginal effects profiles ISCED 2m

Marginal effect and confidence interval
ISCED 2m

<table>
<thead>
<tr>
<th>Marginal effect ISCED 2m</th>
<th>Lower bound (95%)</th>
<th>Upper bound (95%)</th>
</tr>
</thead>
</table>

Age

0 5 10 15 20 25 30

-4.0  -3.0  -2.0  -1.0  0.0  1.0  2.0  3.0  4.0
Results II – Marginal effects profiles ISCED 3m
Results II – Marginal effects profiles ISCED 5m

Marginal effect and confidence interval
ISCED 5m

<table>
<thead>
<tr>
<th>Marginal effect ISCED 5m</th>
<th>Lower bound (95%)</th>
<th>Upper bound (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Age
Results II – Marginal effects profiles ISCED 2f

Marginal effect and confidence interval
ISCED 2f

Age

- Marginal effect ISCED 2f
- Lower bound (95%)
- Upper bound (95%)
Results II – Marginal effects profiles ISCED 3f

Marginal effect and confidence interval
ISCED 3f

Age

Marginal effect ISCED 3f  Lower bound (95%)  Upper bound (95%)
Results II – Marginal effects profiles ISCED 5f

Marginal effect and confidence interval
ISCED 5f

Age

Marginal effect ISCED 5f  Lower bound (95%)  Upper bound (95%)
Results III – Persistence of cohort size effects

• Any significant negative cohort size effects are only temporary
  – ISCED 2m: 12 years (age 26-37)
  – ISCED 3m: 9 years (age 27-35)
  – ISCED 5m: 13 years (age 28-40)
  – [ISCED 5f: 14 years (age 27-40)]

• Future research:
  – Why do marginal effects turn positive?
    • Investment into skill acquisition during period of depressed wages leading to faster wage growth?
    • Move out of perfect competition framework?
Results IV – Comparison across education

• Relative to ISCED 3 the negative effects of cohort size are larger for more educated individuals (males and females):
  – Longer period of predicted negative marginal effects
  – Marginal effects are more negative (profiles are shifted down)

• Results compatible with theory that less educated individuals are better substitutable across age
Results V – Comparison with other studies

• Wright (1991):
  – ISCED 2:
    • Marginal effects: 0.175 (average), -0.153 (minimum); 6 years
  – ISCED 3+5:
    • Marginal effects: 0.004 (average), -0.351 (minimum); 13 years

• Brunello (2010):
  – ISCED 3:
    • Marginal effect: -0.07; constant over age
  – ISCED 5:
    • Marginal effect: -0.175; constant over age

• Own results:
  – ISCED 2:
    • Marginal effects: -0.175 (average), -0.351 (minimum); 12 years
  – ISCED 3:
    • Marginal effects: 0.027 (average), -0.374 (minimum); 9 years
  – ISCED 5:
    • Marginal effects: -0.530 (average); -1.985 (minimum); 13 years
Summary

• Analysis of cohort size-effects on individual wages

• Identification of causal effect using instrumental variables estimation

• Evidence of negative marginal cohort size effects for male wages

• Impact on female wages less clear (labour force participation)

• Negative effects on wages only temporary

• Evidence that negative effects are larger for more educated individuals

• Estimated cohort size effects large in comparison with similar research results