The second Biennial Meeting of the

EARLI Special Interest Group 18 Educational Effectiveness

Leuven, 25-27 August 2010

Paper session 2 (Thursday 9h00 - 10h30)

An analysis of the careers of Italian PhD graduates: are they over-educated?

Matilde Bini & **Leonardo Grilli**

Dipartimento di Statistica "Giuseppe Parenti"

Università degli Studi di Firenze

Outline

- Over-education
- Occupational status of PhD graduates
- Our survey on PhD graduates in Italy
- Analysis via random intercept logit model
- Concluding remarks

Over-education

- **Over-education** (or over-qualification):

  the level of education attained by an individual is *greater than* the education required in the job

- Ways of defining and measuring **required education**:
  - ***Job Analysis*** (JA): systematic evaluation by professional job analysts
  - ***Worker self-Assessment*** (WA): the worker specifies the education required for the job
  - ***Realized Matches*** (RM): required education is derived from what workers in the respondent's job usually have attained (e.g. mean of the distribution of the years of education)

Literature on over-education

- First paper: Duncan & Hoffman (1981) Econ. Educ. Rev.
- Review: Hartog (2000) Econ. Educ. Rev.

Social Sciences Citation Index (ISI)

Topic "overeducation"

2005 to present

89 documents

Field: Subject Area	Record Count	% of 89	Bar Chart
ECONOMICS	54	60.6742 %	
EDUCATION & EDUCATIONAL RESEARCH	17	19.1011 %	
SOCIOLOGY	13	14.6067 %	
INDUSTRIAL RELATIONS & LABOR	7	7.8652 %	
DEMOGRAPHY	4	4.4944 %	
STATISTICS & PROBABILITY	4	4.4944 %	
MANAGEMENT	3	3.3708 %	
SOCIAL SCIENCES, INTERDISCIPLINARY	3	3.3708 %	
SOCIAL SCIENCES, MATHEMATICAL METHODS	3	3.3708 %	
ENVIRONMENTAL STUDIES	2	2.2472 %	
Field: Subject Area	Record Count	% of 89	Bar Chart

(11 Subject Area value(s) outside display options.)

Aims of studies on over-education

- Estimate the incidence
  - in most surveys, 15% to 35% of workers are over-educated
  - Usually, the higher the qualification the higher the risk of over-education (for graduates the incidence is about 30%)
  - the incidence did not increase in the early 1990's (despite the general increase in the educational attainment)
- Evaluate the consequences on
  - Earnings
  - Satisfaction
  - Productivity

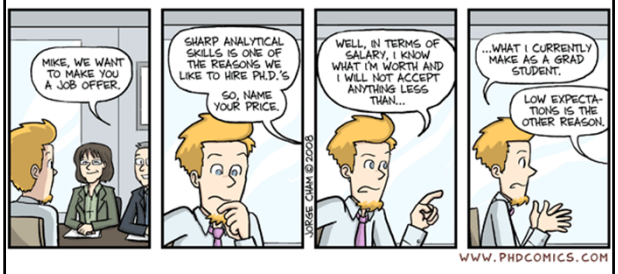
  } Few analyses

Occupational status of PhD's

- The PhD is the highest level of the educational system
- The traditional aim of PhD is to prepare for research (in academic world or in firms)
- During the last decade there has been a rapid growth in the number of PhD graduates
- … but the need of researchers did not increase at the same rate
- Nevertheless, most PhD's are employed, but:

Are their jobs consistent with their education?
Is there a problem of **over-education**?

Often the salaries of PhD's are not excellent …



Here we do not consider salaries but focus on the (self-assessed) consistency between PhD education and the current job

Literature on careers of PhD graduates

- Few studies compared to the careers of graduates with a university degree
- Some relevant studies
  - Martinelli (1999) France
  - Nerad and Cerny (1999) USA
  - Enders (2002) Germany
  - Auriol (2007) seven OECD countries
  - Western, Boreham, Kubler, Laffan, Western, Lawson, Clague (2007) Australia
  - Raddon and Sung (2009) UK
  - … we are working on Italian PhD's

PhD in Italy

- In Italy the PhD was established only in late 80's
  - organized in **annual cycles** and most of them have an institutional length of **three years**
  - about 1/4 of students admitted without grant
- The number of PhD graduates raised dramatically (but now it is decreasing):
  - Year: 1998 2003 2006 2008
  - # grad: 2803 6249 10057 9603
- The demand of PhD graduates comes almost only from universities and a few public research institutions
- The private sector absorbs few PhD graduates and often without requiring the PhD qualification
- The recruitment by universities is slowing down → increasingly, PhD students search for a job outside the research fields → more and more PhD graduates are employed in jobs not requiring their qualifications (over-education)

Our survey on PhD graduates

- Data collected by the University of Florence in 2010 for the National Committee on the Evaluation of the University System
- Population: PhD holders who got their degree from an Italian university in years 1998, 2003, 2008
- The survey intended to reach all PhD's via email or telephone
- For cohorts 1998 and 2003 there is no reliable contact list

  → contact rate is very low

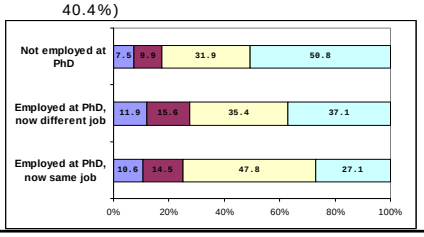
  → ***we consider the 2008 cohort***
- Survey technique: **web questionnaire** + telephone interview for those not filling in the web questionnaire
- Most responses from May to June 2010

The cohort 2008

- Valid responses: 3397 graduates (35% of all graduates, almost all missing graduates are due to failure in contact)
- The employment rate is high: 85.3% are working at the interview
- For the analysis of over-education we consider **2709 employed graduates** (after deleting a few records due to missing values in relevant variables)
- A predictor of over-education is the employment status at PhD graduation for those currently employed:
  - **NOT EMPLOYED AT PHD**: 1353 (49.9%)
  - **EMPLOYED AT PHD, NOW DIFFERENT JOB**: 404 (14.9%)
  - **EMPLOYED AT PHD, NOW SAME JOB**: 952 (35.1%)

Our measure of over-education

- To study over-education, we use the following question *"How much useful is for your job the education acquired during your PhD studies?"*
  - 1: no use (250 9.2%) 2: limited use (335 12.4%)
  - 3: ***useful*** for my approach to work, even if I don't use it in a specific way (1029 38.0%)
  - 4: ***fundamental*** for the tasks I carry out in my job (1095 40.4%)



Category 3 counterbalance Category 4 → don't aggregate those categories!

Areas of PhD courses

- We classify the PhD courses into **7 areas** on the basis of the 14 Italian scientific areas:
  - A. Mathematics, Informatics, Physics (212 7.8%)
  - B. Chemistry, Geology (262 9.7%)
  - C. Biology, Medicine, Veterinary, Agriculture (846 31.2%)
  - D. Engineering, Architecture (483 17.8%)
  - E. Literature, Arts, Philosophy, History, Pedagogy, Psychology (476 17.6%)
  - F. Law (155 5.7%)
  - G. Economics, Statistics, Social Sciences, Political Sciences (275 10.1%)

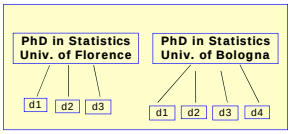
area	%empl	%same	%fundam
A	92	24	53
B	88	29	42
C	85	39	39
D	88	33	41
E	80	38	36
F	80	45	39
G	89	32	43

Multilevel structure

- The 2709 employed graduates are clustered in 1251 PhD courses

size	Freq.	Percent
1	567	45.3
2	335	26.8
3	167	13.4
4	82	6.6
5+	100	8.0

- Level 2: PhD course in a given university
- Level 1: PhD graduates



We need a model with **random effects** to account for the unobserved heterogeneity among PhD courses

Random effects proportional odds model

- graduate i in PhD course j
- categories of the ordinal response $1, 2, \ldots, C$
- cumulative probabilities $\gamma_{ij}^{(c)} = P(Y_{ij} \le c \mid \mathbf{x}_{ij}, \mathbf{w}_j, u_j)$
- *proportional odds* model

$$\log\left(\frac{\gamma_{ij}^{(c)}}{1-\gamma_{ij}^{(c)}}\right) = \alpha_c - \left(\boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + u_j\right) \qquad c = 1, \ldots, C-1$$

with α_c cut-points
u_j random effect for PhD course j

Response = "How much useful is for your job the education acquired during your PhD studies?" → *C*=4 categories

In the model without covariates
- estimated cluster variance =0.154 (ICC=0.047 on the latent scale)
- p-value of the LRT test on the cluster variance = 0.013

Problems with proportional odds

- The proportional odds model fully exploits the ordinal nature of the response … at the cost of a strong assumption
- Indeed the Brant test (Biometrics 1990) rejects the ***proportional odds assumption*** for the binary covariate "EMPLOYED AT PHD, NOW SAME JOB"
- Solutions:
  1. Exclude the graduates with same job
  2. Extend the model to relax the assumption (partial proportional odds model)
  3. Collapse the first three categories of the response and use a logit model for the probability that PhD is fundamental for the current job
- We choose solution #3: it gives insight into the issue with a simple model

Random effects logit model

- graduate i (level 1) in PhD course j (level 2)
- response Y_{ij}=1 ↔ the education acquired during PhD is ***fundamental*** for the current job
- conditional probability $p_{ij} = P(Y_{ij} = 1 \mid \mathbf{x}_{ij}, \mathbf{w}_j, u_j)$

$$\text{logit}\,(p_{ij}) = \alpha + \boldsymbol{\beta}\mathbf{x}_{ij} + \boldsymbol{\gamma}\mathbf{w}_j + u_j$$

with u_j random effect for PhD course j

Strategy for model selection:
- Null model (sd=0.478, ICC=0.065)
- Model with level 1 covariates (sd=0.414, ICC=0.049)
- Model with level 1 + level 2 covariates (sd=0.275, ICC=0.023)

In parenthesis the std.dev. of the random effects and the ICC on the latent scale

The addition of the level 2 covariates reduces substantially the std.dev. of the random effects (which is no more statistically significant) → the unobserved heterogeneity among PhD courses is modest

Estimates

Random intercept logit model for the probability that the education acquired during PhD is fundamental for the current job

LEVEL 1 COVARIATES	
Not employed at PhD	ref.cat.
Employed at PhD, now different job	-0.32 *
Employed at PhD, now same job	-0.79 *
Unrelated job during PhD	-0.57 *
No grant for PhD	-0.47 *
Period abroad during PhD	0.62 *
LEVEL 2 COVARIATES	
Rating: content of courses	0.00
Rating: quality of teaching	-0.07
Rating: scientific level of researchers	0.09
Rating: training to do research	0.18 *
Rating: opportunities to publish	0.08 *
Math, Physics	ref.cat.
Chemistry, Geology	-0.31
Biology, Medicine, Agriculture	-0.25
Engineering, Architecture	-0.22
Literature, Arts, Philosophy, Psychology	-0.24
Law	-0.15
Economics, Social and Political Sciences	0.12
(intercept)	-1.83 *
CLUSTER LEVEL STANDARD DEVIATION	0.28

2705 graduates in 1247 PhD courses

*Significant at 5%

Ratings (1 to 10) averaged within each PhD course

Area of PhD course

Main findings

- The usefulness of the PhD education for the current job is strongly related with the condition of the student during PhD
  - The higher chances that the PhD will be useful is for conventional "good" students, namely those who do not work at all, have a grant, spend a period abroad
- The statistically significant features of the PhD course are the ability to train for research and to give opportunities to publish
  - The content and quality of teaching are not significant
- The covariates explain most of the differences among PhD courses: the std.dev. of the random effects and the dummies for the areas are not significant
  - The higher chances for PhD graduates in Math/Physics (53% vs 40% overall) are explained by good values of the covariates

Distribution of covariates in areas

Graduate-level covariates AREA	PhD fundamental (%)	Employed at PhD, now same job (%)	Unrelated job during PhD (%)	No grant for PhD (%)	Period abroad during PhD (%)
Math, Physics	**53**	24	16	18	48
Chemistry, Geology	42	29	16	28	39
Biology, Medicine, Agriculture	39	39	19	33	32
Engineering, Architecture	41	33	24	27	39
Literature, Arts, Philosophy, Psychology	36	38	**34**	**38**	45
Law	39	**45**	31	30	43
Economics, Social and Political Sciences	43	33	28	30	**57**

PhD-level covariates AREA	Rating: training to do research	Rating: opportunities to publish
Math, Physics	**7.1**	**7.2**
Chemistry, Geology	**7.1**	6.5
Biology, Medicine, Agriculture	6.8	6.4
Engineering, Architecture	6.6	6.4
Literature, Arts, Philosophy, Psychology	6.3	5.5
Law	6.6	6.1
Economics, Social and Political Sciences	5.8	5.2

Alternative merging of the categories

- Response *"How much useful is for your job the education acquired during your PhD studies?"*
  - 1: no use 2: limited use
  - 3: ***useful*** 4: ***fundamental***
- We dichotomised 4 vs (1+2+3)
- What happens if we dichotomize (3+4) vs (1+2)?
  - It happens that the effects of the covariates are attenuated, e.g.
    - Employed at PhD, now same job: -0.79 becomes -0.20
    - Period abroad during PhD: 0.62 becomes 0.46
    - Opportunities to publish: 0.08 becomes 0.05
  - The variance of the random effects is negligible

The (3+4) merging makes PhD graduates and PhD courses alike → it reduces the discriminating power of the covariates

Email: grilli@ds.unifi.it

Web: www.ds.unifi.it/grilli