

# Linking NPD and LEO



Dave Thomson

Education Datalab

Tw: @edudatalab

[www.educationdatalab.org.uk](http://www.educationdatalab.org.uk)

# Longitudinal Educational Outcomes (LEO)



- Administrative DWP/ HMRC data matched to NPD
- Created by DWP
- Mostly consists of dates on which events occurred
  - Benefit histories (from 1999/2000)
  - PAYE Employment (P45) data (from 1999/2000)
  - Earnings from employment data (from 2004/05)
- In principle, we can observe whether an individual on any given day
  - Was in employment
  - Was in receipt of benefits (**but not yet universal credit**)



# LEO- data limitations

- Data received from DfE has been cleaned/  
transformed from raw data supplied by DWP
  - Missing (or default) start and end dates in  
employment data
  - End dates for benefits not measured precisely
- Other missing data
  - No universal credit data
  - No 'earnings' (dividends) for the self-employed
  - In the past, earnings did not have to be reported  
for those earning below the lower tax threshold



# Longitudinal Educational Outcomes (LEO)



- Used by IFS to calculate relative labour market returns to degrees/ CVER returns to vocational education
- Also used by DfE in calculation of destinations measures for Performance Tables
- No process currently in place (yet) to share with researchers beyond those doing research commissioned by DfE
- Currently working on DfE-sponsored project looking at pathways through the education system with Universities of Westminster and Bath



# Pathways through the education system



- 8 cohorts of young people who turned 16 between 2001/02 and 2008/09
- Data from NPD (KS4, census etc.) has been linked to FE data (ILR), HE data (HESA), employment, earnings and benefits (LEO)



# Attrition



Age last observed in any source (NPD, ILR, HESA, LEO)

KS4 Year	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	People
2002	98%	97%	95%	94%	93%	92%	91%	91%	90%	89%	88%	86%	85%	83%	80%	588214
2003	97%	96%	94%	93%	93%	92%	91%	90%	89%	88%	86%	85%	83%	79%		639882
2004	98%	97%	95%	94%	94%	93%	92%	91%	90%	88%	86%	84%	81%			660947
2005	98%	97%	96%	95%	95%	94%	93%	91%	90%	88%	86%	82%				639611
2006	99%	98%	97%	96%	96%	95%	93%	92%	90%	88%	84%					649189
2007	99%	98%	97%	97%	96%	95%	93%	92%	89%	85%						658189
2008	99%	98%	98%	97%	96%	95%	93%	91%	86%							656404
2009	98%	98%	97%	97%	96%	94%	92%	86%								648341



# % with observed earnings by age

KS4 Year	% with observed earnings																People
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
2002		37%	56%	61%	61%	62%	64%	68%	72%	73%	71%	72%	72%	72%	71%	588214	
2003	28%	45%	58%	59%	59%	61%	65%	70%	72%	70%	71%	71%	71%	71%		639882	
2004	33%	47%	58%	59%	60%	64%	70%	73%	72%	73%	73%	73%	72%			660947	
2005	33%	47%	58%	60%	62%	69%	73%	73%	75%	75%	75%	75%				639611	
2006	32%	46%	58%	61%	67%	72%	73%	76%	77%	77%	76%					649189	
2007	31%	45%	57%	65%	69%	71%	75%	77%	78%	77%						658189	
2008	27%	42%	59%	68%	68%	73%	77%	79%	78%							656404	
2009	25%	43%	61%	65%	70%	75%	78%	78%								648341	



# Qualification Choices and Routes

- Can use the data to look at the long-term outcomes associated with different qualification choices at a very detailed level
- An example: the General National Vocational Qualification (GNVQ)





# What's a GNVQ?

- Introduced in the early 1990s to provide experiences and education relevant to working life but stopping short of providing training for a specific job.
- They could be studied at foundation level (equivalent to GCSE grades D-G), intermediate level (equivalent to GCSE grades A\*-C) and advanced level (equivalent to A-level) and were available in 15 broad vocational areas.
- The full GNVQ consisted of six modules, each of which had a recommended guided learning time of 60 hours. A part 1 GNVQ consisted of three modules.
- Two-thirds of the modules were internally assessed based on coursework. The full GNVQ was considered equivalent to 4 GCSEs and the part 1 GNVQ equivalent to 2 GCSEs.



# GNVQ and Performance Tables



Support The Guardian

Subscribe Find a job [Sign in / Register](#) Search ▾

UK edition ▾

**News**

**Opinion**

**Sport**

**Culture**

**Lifestyle**

More ▾

# The Guardian

Education ▶ Teachers

**Schools**

## Couldn't do better

**Can Thomas Telford notch up another 100% GCSE success? asks Phil Revell**

**Phil Revell**

Tue 21 Aug 2001 12.01 BST



27

All eyes will be on the GCSE results on Thursday, with special attention being paid to one technology college in Shropshire. Can Thomas Telford's students repeat last year's remarkable 100% success rate in the exam?

Advertisement



FOR EVERY CHILD IN DANGER

RECEIVE FREE RESOURCES FOR YOUR CLASSROOM

SIGN UP



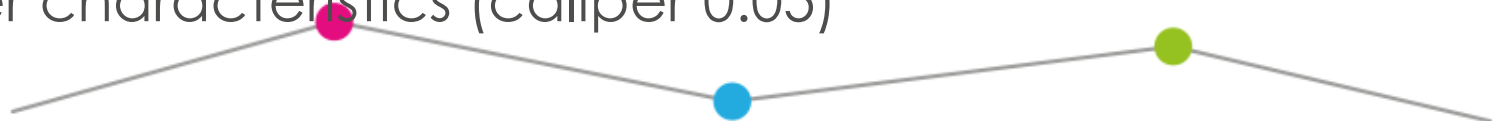
# The unresolved dilemma of equivalence

		Grade	Points	Size (GCSE equivalents)	Equivalent to GCSE grade
GNVQ Full	Intermediate	D	30	4	A*-A
		M	24	4	B
		P	20	4	C
		U	0	4	U
	Foundation	D	16	4	D
		M	12	4	E
		P	6	4	F-G
		U	0	4	U
GNVQ Part 1	Intermediate	D	15	2	A*-A
		M	12	2	B
		P	10	2	C
		U	0	2	U
	Foundation	D	8	2	D
		M	6	2	E
		P	3	2	F-G
		U	0	2	U



# Long-term outcomes of a GNVQ at 16

- 544 thousand pupils in 2002/03 cohort of 16 year olds at state-funded mainstream schools
  - Treatment pool: 121 thousand pupils entered for at least 1 GNVQ
  - Comparison pool: 169 thousand pupils at schools which did not enter any pupils for GNVQ
- Comparison groups
  - Calculated separately for each of the 4 types of GNVQ
  - Exact matching on region and free school meal eligibility
  - Propensity score (nearest neighbour) matching on other characteristics (caliper 0.05)



# Matching

- 121 thousand in the treatment groups
- 6 thousand lost due to missing data (mostly KS2)
- 5 thousand unmatched due to caliper constraint
- 110 thousand matched



# Balancing



		Before Matching		After Matching	
		GNVQ	Not GNVQ	GNVQ	Not GNVQ
Government Office Region	North East	7%	4%	7%	7%
	North West	15%	16%	15%	15%
	Yorkshire & Humber	9%	12%	9%	9%
	West Midlands	16%	8%	14%	14%
	East Midlands	8%	8%	8%	8%
	East of England	12%	14%	12%	12%
	Inner London	3%	3%	3%	3%
	Outer London	9%	10%	9%	9%
	South East	12%	16%	12%	12%
	South West	9%	9%	10%	10%
Term of birth	Autumn	24%	24%	24%	24%
	Winter	24%	24%	24%	24%
	Spring	26%	26%	26%	26%
	Summer	27%	27%	27%	27%
Pupil Characteristics	Female	48%	50%	48%	48%
	FSM	15%	11%	15%	15%
	SEN: School Action	9%	8%	9%	9%
	SEN: School Action Plus or Statement	5%	6%	5%	5%
	EAL	10%	7%	9%	9%
Pupil prior attainment (Key Stage 2)	English	4.16	4.27	4.16	4.16
	Maths	4.11	4.22	4.11	4.1
	Science	4.25	4.34	4.25	4.24
Pupil prior attainment (Key Stage 3)	English	5.31	5.51	5.31	5.3
	Maths	5.57	5.79	5.57	5.56
	Science	5.37	5.57	5.38	5.37
School prior attainment	KS2	4.15	4.27	4.17	4.17
	KS3	5.35	5.59	5.38	5.38



# Outcomes



## Outcome

## Source

Achievement of NQF level 3 (2 A-level or equivalent) by age 19

NPD Post-16 achievement data

Achievement of 2 A-levels by age 19

NPD Post-16 achievement data

Higher Education Participation by age 22

HESA Student Record

Employed for at least 180 days in 2015/16

Matched HMRC employment (P45) data

Earnings in 2015/16

Matched HMRC earnings (P14) data



# Results



GNVQ Type	Group	Level 3 by age 19	2 A- levels by age 19	HE participation by age 22	Employed age 29	Earnings age 29	Earnings Included	Attrition	Number
Part 1 Foundation	GNVQ	13%	4%	6%	62%	£15,345	62%	12%	13017
	Not GNVQ	17%	6%	10%	62%	£15,806	62%	12%	13017
Full Foundation	GNVQ	13%	3%	6%	62%	£15,623	62%	12%	3593
	Not GNVQ	17%	7%	10%	61%	£16,053	62%	12%	3593
Part 1 Intermediate	GNVQ	49%	32%	33%	73%	£20,483	72%	10%	55126
	Not GNVQ	49%	34%	35%	71%	£20,479	71%	10%	55126
Full Intermediate	GNVQ	50%	34%	34%	74%	£21,247	73%	10%	38474
	Not GNVQ	48%	34%	35%	71%	£20,832	71%	10%	38474
All	GNVQ	44%	29%	29%	71%	£20,088	71%	11%	110210
	Not GNVQ	44%	30%	31%	70%	£19,980	70%	11%	110210





# Summing up

- LEO linked to NPD potentially useful in quantifying the returns to different qualifications
- Could offer a perspective on questions of the equivalence of qualifications taken by young people
- There are holes in the LEO data

