Evaluating Projections of Changing Skill Needs

International Symposium on Skills Forecasting
University of Warwick, September 2011
Rob Wilson

Professorial Fellow, Institute for Employment Research, University of Warwick, United Kingdom

Overview

General considerations

- ➢ Background - rationale for skills projections
- ➢ Quantitative metrics
- ➢ Causes of errors

Alternative approaches to evaluating projections

- ➢ *Quantitative* evaluation of the results (internal/external); and
- ➢ *Qualitative assessments*

  (getting the message across; review of users opinions)

Previous attempts to assess skills projections

Some tentative conclusions

General Considerations in Evaluating Labour Market Forecasts

Accuracy, value and usefulness - Should we expect accuracy in labour market forecasting?

Understanding the rationale for employment forecasting - What is the purpose of forecasts?

➢Making technical information digestible for a lay audience

➢To guide individual decisions and highlight alternatives

➢Setting out options to policy makers

➢To influence and change behaviour, including policy

Therefore outcomes may be EXPECTED to be different from forecasts

Need for a qualitative as well as quantitative assessment of employment forecasts

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Forecasts can:

- Make assumptions about the Future **Explicit** & **Transparent**
- Help to Enforce **Systematic** and **Logical** Thinking
- Act as a **Focus** for intelligent **Debate**
- Provide a useful **counterfactual** to assess policy impacts
- Explain past developments in a coherent and logical fashion and then how this behaviour may develop into the future

But they cannot deliver:

- Mechanistic "manpower" plans
- Precise indications of education and training requirements or job opportunities
- Not a crystal ball: impossible to foresee the unforeseen

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Quantitative and qualitative assessment

- Quantitative measures of forecast accuracy
- Metrics, Losses and Scores
  - ➢ Absolute (000s) and relative (%) errors
  - ➢ Direction of change
  - ➢ Significance / importance

- Qualitative assessments of:
  - ➢ accuracy
  - ➢ usefulness

Quantitative Metrics

Absolute error = $Z_i - \check{Z}_i$ *(may be positive or negative)*

Relative error = $(Z_i - \check{Z}_i) / Y_i$

   Z_i represents the outcome for some particular quantity being projected (e.g. Employment in a particular occupation and/or qualification category), $\check{Z}_i$ represents the projected value and Y_i indicates the level of employment in that category i. (by definition $Z_i = Y_i$ for expansion demands)

Mean absolute percentage error

   (Absolute value of Absolute error; weighted, MAPE(wt) and unweighted, MAPE)

Direction of change (DoC)

Rank by change (RbC

   Spearman rank correlation coefficients (absolute & percentage change)

Dissimilarity Index (DI)

   (focuses on employment shares)

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Loss, Average Loss & Scores

Loss: $L_i (\check{Z}_i) = ((Z_i - \check{Z}_i) / Y_i)^2$

Average loss: (AL) is determined by weighting the separate losses according to the size of the occupational or qualification category

Forecasts should be compared to some plausible alternative forecast method not some unattainable ideal of "perfect accuracy"

The most basic alternative is *SAB:* (Same as Before) which assumes no changes between the base and the forecast year – various other "naïve" alternatives can be used (past trends, etc)

Score: the ratio between the predictive quality of the forecast (loss or average loss) and the predictive quality of some such alternative reference forecast (*Score* is smaller than 1 if the forecast is better than the reference forecast and is more than 1 if the forecast is inferior to the reference)

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH   IER

Qualitative evaluation

Qualitative evaluation of projections (illustrative)

Outcome/ Forecast	Fast decline	Low decline	Little change	Low growth	High growth	all
Fast decline	5	5	0	0	0	10
Slow decline	6	10	4	0	0	20
Little change	4	0	10	4	2	20
Slow growth	0	0	10	6	4	20
Fast growth	0	0	1	0	10	11
Total	15	15	25	10	16	81

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Explaining the Errors - Problems in measurement & possible source of data errors

History

➢ Historical data revisions

➢ Changes in classification

➢ Error in the measurement of the **dependent variable**

➢ Errors in the **endogenous drivers**

and *Forecasts*

➢ **Errors of judgement** about exogenous variables;

➢ **Errors of specification,** reflected in the various **parameters** of the model;

➢ **Errors of judgement** in running the models

Distinction between Scale (macro), Sectoral & Occupational effects

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Previous and recent attempts to assess skills forecasts

US BLS projections (Bishop *et al*, Rosenthal and most recently Wyatt)

Dutch experiences (ROA, Borghans and deGrip, Dupuy)

Australia (Meagher *et al*)

UK (Wilson, Wilson *et al*)

Pan- European / Cedefop (Kriechel, Pollitt and Wilson)

Assessments of US Projections

- Broadly correct, but errors in detail
- Industries
  - ➢ BLS does better than the naïve models in all metrics:
    - ❖ *DoC*
    - ❖ *MAPE & MAPE(wt)*
    - ❖ *DI*
    - ❖ *RbC (both Spearman RCCs)*
- Occupations
  - ➢ BLS does better than the naïve models in 3 of 4 metrics:
    - ❖ *DI*
    - ❖ *MAPE & MAPE(wt)*
- *Generally better at projecting larger occupations*

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Summary of UK experience

- Broadly correct, but errors in detail
- Industries
  - Margins of error similar to average changes in the variables
  - Much of the error attributable to revisions to historical data or to shifts in exogenous factors
- Occupations
  - Margins of error slightly less than average changes in the variables (but uses a higher level of aggregation than BLS)
  - Replacement demands smaller % errors

The Dutch experience

- Covers Qualifications as well as Occupations, Demand and Supply
- Broadly correct in qualitative terms
- Replacement demands better "*Loss*" than for Expansion demands
- Generally average *Loss* and *Scores* improving over time
- Changes in industry classification caused considerable problems recently
- But many *Scores* in excess of unity especially for Replacement Demands suggests further room for model improvements

Cedefop projections

- Evaluation a key component of the project
- Quantitative and qualitative aspects
- Qualitative:
  - ➢ Report on survey held among SkillsNet members and AGORA 2009 & 2010 visitors
  - ➢ Response (n=65); Of which (n=48 "experts")
  - ➢ Topics include:
    - ❖ Use of results
    - ❖ Evaluation of outcomes

Assessing the Cedefop Pilot projections – Quantitative results

Very early days, pilot projects only undertaken in 2008
 and very short times series (projections made only shortly before the GFC)

By far the largest difference in both actual outcomes, due to the economic crisis

Makes any short-term assessment difficult - for example, productivity changes in 2008-09 would have been impossible to predict

But apart from levels, other trends not so dramatically affected (some errors by industry, less significant by occupation and qualification)

Qualitative evaluation – General approach

Usage and Usability of CEDEFOP Forecasting Model

Qualitative Evaluation: Focus groups

Qualitative evaluation: Internet Survey Instrument

Qualitative Evaluation -Topics

I. **Usability**
  i. **Workbooks**
  ii. **Actual usage**
  iii. **Interactive**

II. **Outcomes**
  i. **Sector**
  ii. **Occupation**
  iii. **Education**
  iv. **Supply / Demand**
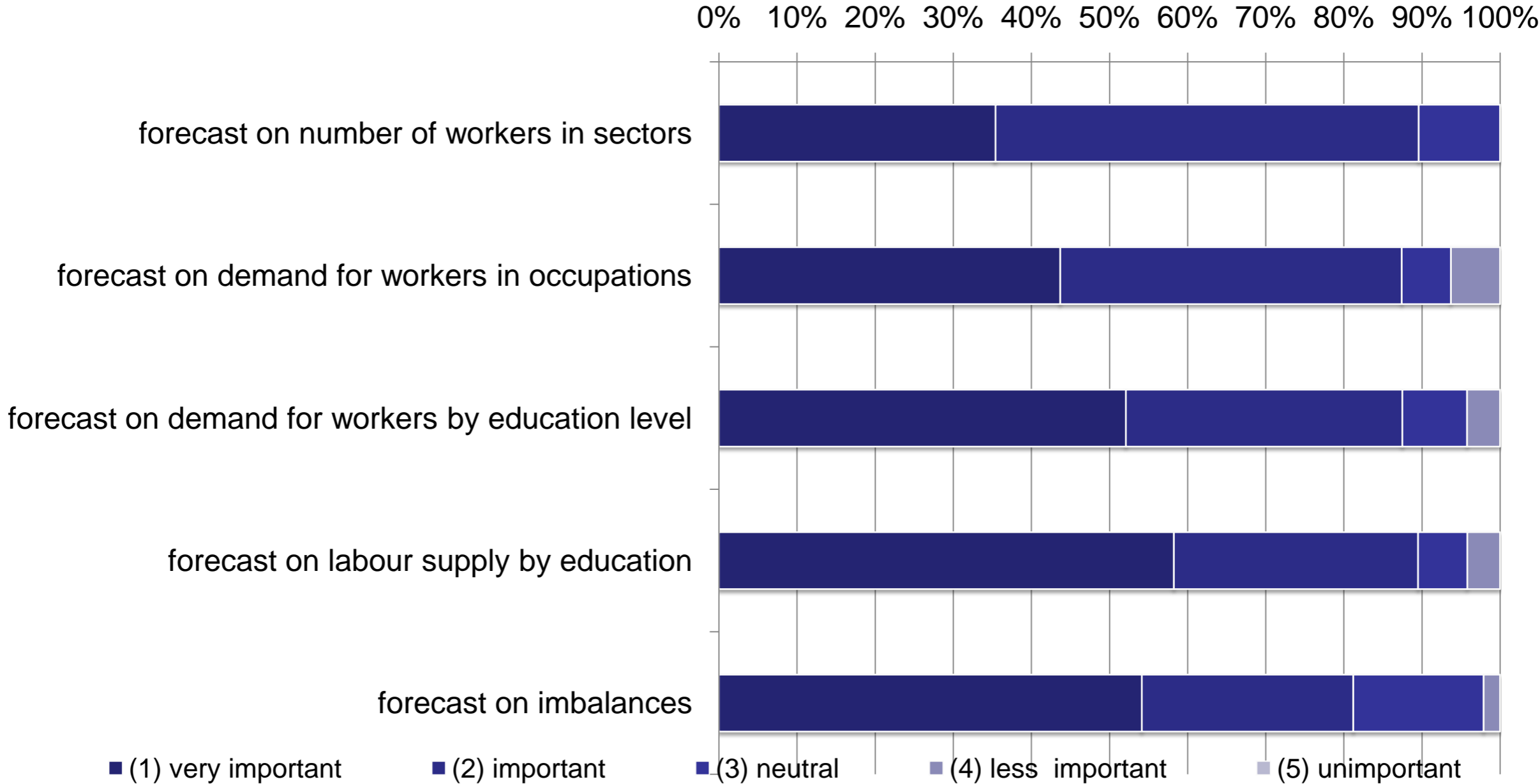 v. **Level of aggregation**

III. **Methodology**
  i. **Sector model**
  ii. **Scenarios**
  iii. **Expansion demand**
  iv. **Replacement demand**
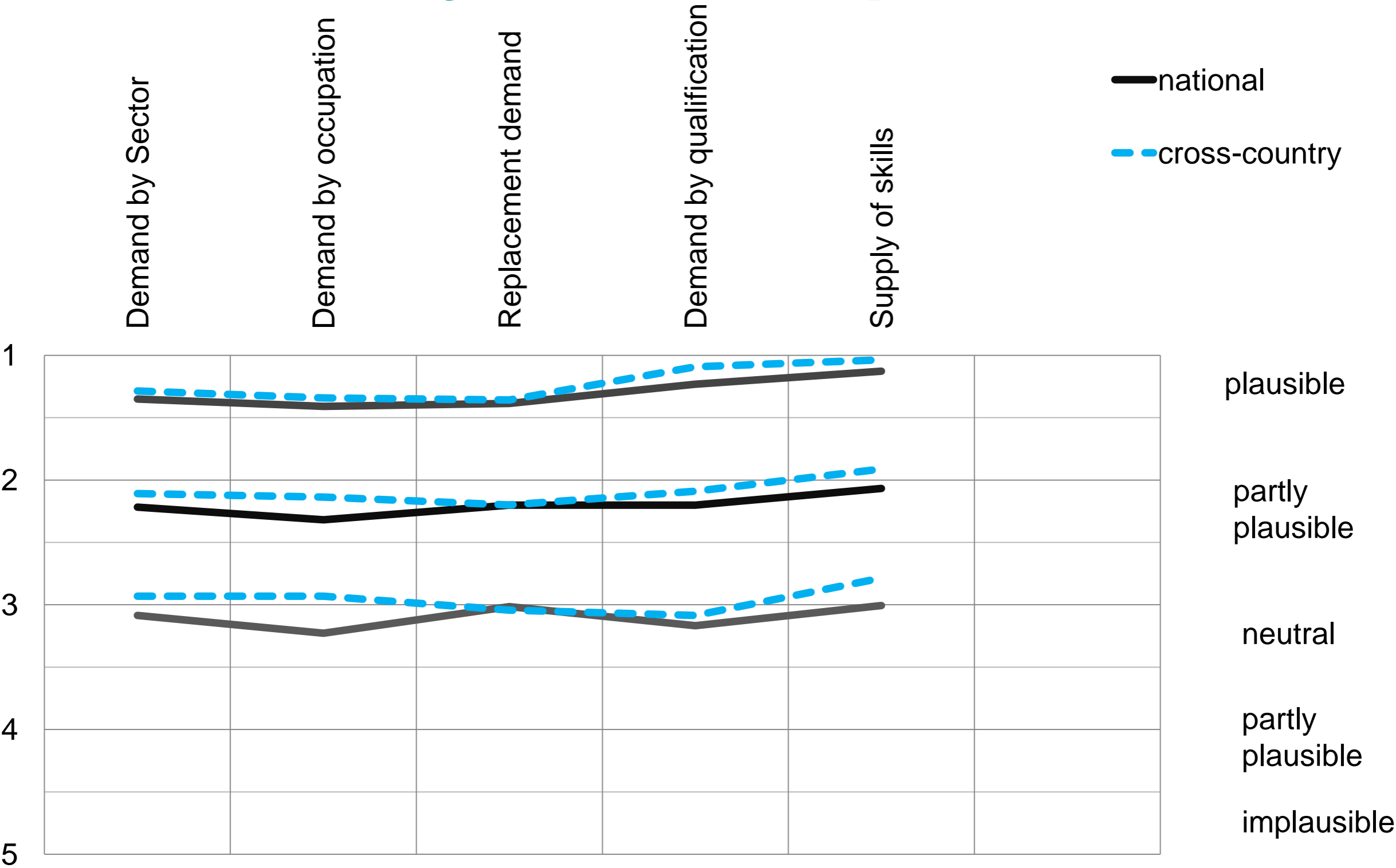 v. **Supply (stock / flow)**

Comparison of overall Cedefop results to national forecasts

	Percent
Don't know	*10,87*
Very close to national forecast / expectations	*6,52*
Same direction of outcomes	*50,00*
Partially contradicting outcomes	*30,43*
Contradicts national forecasts / expectations	*2,17*
n	46

Importance of different aspects of the Cedefop forecasts

Plausibility of Cedefop forecasts

Demand by Sector

Demand by occupation

Replacement demand

Demand by qualification

Supply of skills

national

cross-country

1 — plausible

2 — partly plausible

3 — neutral

4 — partly plausible

5 — implausible

Value added of Cedefop forecast relative to national forecasts

| 22.2 | 33.3 | 31.5 | 9.3 | 3.7 |

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

- (1) very much value added
- (2) much value added
- (3) some value added
- (4) not much value added
- (5) no value added

Value of Workbooks

	Able to change underlying data	Detailed data in workbooks
(1) very important	19,57	52,17
(2) important	47,83	39,13
(3) neutral	21,74	4,35
(4) less important	2,17	4,35
(5) not important	8,7	0
Observations	46	

Some tentative conclusions

- Nobody has a crystal ball - art as much as a science
- Forecasts may be better to be wrong!
- Projections not inevitable, the future is not predetermined - may be able to take actions to improve things
- Skills forecasts should be compared with viable alternatives & generally add value to naïve alternatives
- Historical data revisions are a key source of error, as are political and other exogenous events
- Forecasts are useful and broadly accurate
- Sectoral focus is key
- Methods improving, errors no worse than in other forecasts
- It is important to develop better, more detailed, data, Europe still has some way to go

WARWICK INSTITUTE *for* EMPLOYMENT RESEARCH IER

Contact Details

Rob Wilson

Institute for Employment Research

University of Warwick

COVENTRY, CV4 7AL

r.a.wilson@warwick.ac.uk

Tel:+(44) 2476-523530