European Centre for the Development of Vocational Training

RESEARCH PAPER

No 25

# Skills supply and demand in Europe

## Methodological framework

# Skills supply and demand in Europe

## Methodological framework

Luxembourg: Publications Office of the European Union, 2012

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://europa.eu).

Cataloguing data can be found at the end of this publication.

Luxembourg: Publications Office of the European Union, 2012

ISBN 978-92-896-1112-1
ISSN 1831-5860
doi: 10.2801/85871

The **European Centre for the Development of Vocational Training** (Cedefop) is the European Union's reference centre for vocational education and training. We provide information on and analyses of vocational education and training systems, policies, research and practice. Cedefop was established in 1975 by Council Regulation (EEC) No 337/75.

Europe 123, 570 01 Thessaloniki (Pylea), GREECE
PO Box 22427, 551 02 Thessaloniki, GREECE
Tel. +30 2310490111, Fax +30 2310490020
E-mail: info@cedefop.europa.eu
www.cedefop.europa.eu

# Foreword

Consistently high unemployment rates indicate that present labour market imbalances are also due to structural problems that have been aggravated by the economic crisis. This also makes it more difficult to meet the challenges of several mega-trends, which are changing skill requirements and needs. Demographic developments such as ageing and migration affect the structure of the labour force. ICT applications in most professions and a need to strengthen green skills and sustainable growth have an impact on skill needs.

Adequate policy responses must take account of current and future labour market developments, understand and anticipate future supply and demand for skills.

Cedefop’s skills demand and supply forecast is one tool to understand better and anticipate labour market needs. Forecasts do not offer a crystal ball but signal trends and support the design of long-term policy strategies to mitigate risks.

Cedefop’s forecast does not replace national efforts to anticipate skill needs. On the contrary, it complements such efforts and provides assistance and guidance to national experts.

Cedefop’s forecasting results are based on our efforts to develop a robust and reliable methodology, which is explained in detail in this publication. The detailed description of Cedefop’s forecasting method and the latest accuracy tests demonstrate the quality of our findings but also invite input for further developments.

We hope that this publication will also inspire those who are about to start such a project on the national or regional level and stimulate those who already use similar tools in their search for improvements.

Christian F. Lettmayr
Acting Director

# Acknowledgements

This publication is the result of a team effort and reflects the contributions of all those involved in the project. Alena Zukersteinova and Vladimir Kvetan (Cedefop) ensured overall coordination of the project, under the supervision of Pascaline Descy, head of Research and Policy Analysis. Rob Wilson, Derek Bosworth and Ilias Livanos from the Institute for employment research (IER); Terry Ward and Robert Stehrer from Alphametrics; Ben Gardiner, Hector Pollitt, Unnada Chewpreecha from Cambridge Econometrics and Ben Kriechel, Jan Sauermann and Arnaud Dupuy from the Research Centre for Education and the Labour Market (ROA) form the research team responsible for producing projections and background documentation for Cedefop's forecast (Cedefop contract No 2008-FWC12/AO/RPA/AZU-TODU/European – skills-forecast/009/08).

Our gratitude goes to all Skillsnet members and other experts from European countries who took an active role in the technical workshops. We appreciate their reviews and comments on the methodological developments and key findings.

We would also like to thank Konstantinos Pouliakas for his very useful review of the publication and Roula Panagiotou for her valuable assistance to the project.

This project is supported financially by the 'Progress programme' – a European Union programme to promote employment and social solidarity (2007-13). This programme is managed by the Directorate-General for Employment, Social Affairs and Equal Opportunities of the European Commission.

# Table of contents

# List of tables and figures

## Tables

# Figures

CHAPTER 1.
# Introduction

## 1.1. Background and rationale

Equipping the labour force with the right skills is one of the key policy focuses of the European Union’s (EU) strategy for smart, sustainable and inclusive growth. Anticipation of skill needs has received more attention in the EU, as illustrated by several policy documents such as the *Integrated guidelines for growth and jobs* (2008-10) (European Commission, 2007), the Council Resolution on new skills for new jobs, and the Spring 2008 Council Conclusions (Council of the European Union, 2007; 2008), in which the EU Member States asked the European Commission to report on future skills requirements in Europe up to 2020. *New skills for new jobs: action now* (European Commission, 2010a), a report prepared by an expert group set up by the European Commission, also emphasised the need for a coordinated approach to improve Europe’s capacity to anticipate change. After a wide consultation among stakeholders, the Commission included the New skills and jobs agenda as a flagship initiative in the EU 2020 strategy (European Commission, 2010b) ([1]).

It is in this context that Cedefop conducts regular, coherent and systematic skill demand and supply forecasts. In 2008, Cedefop released the first pan-European skill needs forecast, i.e. employment projections by sector, occupation and qualification level across Europe up to 2015. In 2009, Cedefop forecast the supply of skills by gender, age group and qualification level. Finally, in 2010 the first parallel forecast of skill supply and demand up to 2020 was presented.

A variety of forecasting methods are used. The accuracy of methods must be tested and compared with available alternatives to increase the quality of results. At the same time statistical authorities publish new data regularly. Finally, forecasting is an ongoing exercise, affected by changing reality, which means it is important to use the most up-to-date information and to reflect trends and changes to achieve the most reliable results.

This publication presents the complex methodological framework used by Cedefop to forecast skills supply and demand and some current attempts to improve it. It does not promote Cedefop’s methodology as the only correct methodology. Moreover, Cedefop’s forecast does not replace those conducted at

---

([1]) http://ec.europa.eu/eu2020/.

national level. Instead, this publication presents the problems that we have encountered and the solutions we have adopted to produce a unique pan-European skills supply and demand forecast

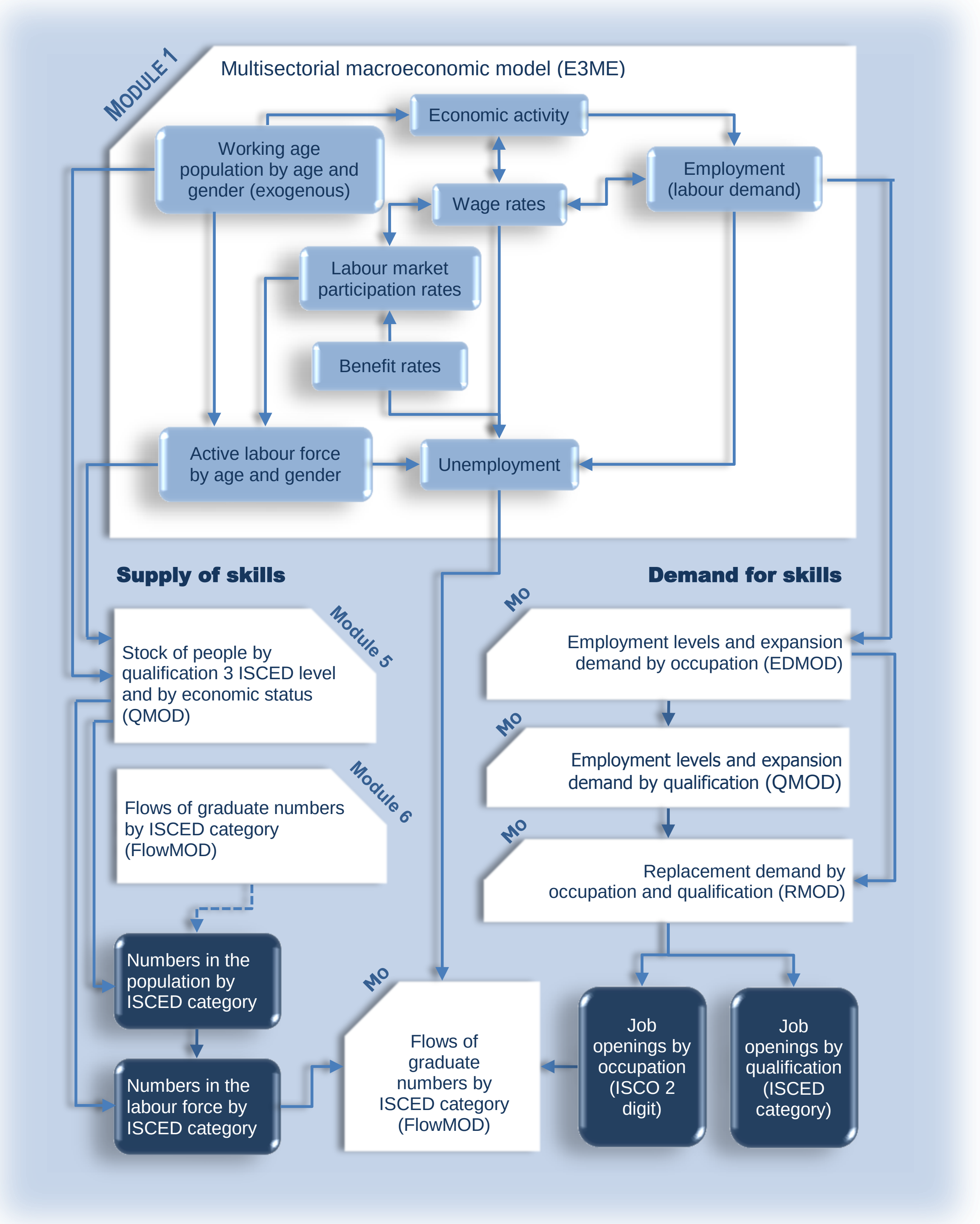
## 1.2. General overview of the methodological framework

Europe's pan-European forecast of skill needs requires complex methods, relying on long-term research and drawing on the expertise of several high-level European research institutions. The modelling tools have been designed to enable further development and customisation. The general framework consists of methods developed in two pilot studies on *Future skills needs in Europe* (Cedefop, 2008) and *Future skill supply in Europe* (Cedefop, 2009). These were combined to produce the first pan-European forecast of skills supply and demand in Europe (Cedefop, 2010). Forecasting is a dynamic process, and important developments took place in 2010 made possible by the modular approach adopted, which enables the different parts of the system to be improved independently. As shown in Figure 1, the model breaks down into different building blocks and into several interrelated components.

Even though the modelling framework has proven to be rather robust, a dialogue must be established with experts from European countries, who are likely to have much greater knowledge of employment trends and data sources within their own countries. By making it easy to incorporate new data and alternative or additional assumptions, the modelling framework provides an opportunity for knowledge and input of experts to be built in efficiently and transparently.

The project involved developing consistent databases and related tools to produce a comprehensive and consistent set of skill projections for all EU Member States plus Norway and Switzerland (EU-27+). The system, models and modules rely upon official data sources, drawing primarily on Eurostat, in particular on Eurostat demographic data, national accounts (NA), the EU labour force survey (EU-LFS), as well as additional data on flows of qualifications. Compilation and harmonisation of the best possible data available for measuring employment was a major achievement of the project. Historically, most countries have invested considerable resources in developing data for their NA. In many respects estimates of employment on this basis are to be preferred as they are consistent with other key economic indicators, such as output and productivity. On the other hand, the EU-LFS has the advantage of providing measures of employment structured by skills (occupation and qualification), as well as by gender and age, which are not available from NA-based estimates.

Figure 1. **Interaction within the forecasting system**




*Source*: Cedefop (2010).

### 1.2.1. Supply of skills

The skill supply projections produce consistent pan-European projections broken down by age, gender and formal qualification ([2]). The results indicate the future skill supply by highest qualification held as well as by age groups and gender for the population and labour force aged 15 and over. The skill supply projections are compatible with the skills demand projections (when focusing on qualifications).

The historical analysis and projections of overall labour supply by age and gender are provided by an extended version of the existing pan-European macroeconomic model E3ME developed by Cambridge econometrics ([3]), which incorporates a new demographic and labour-supply module. E3ME models labour supply as a function of economic activity, real wage rates, unemployment and other benefit rates. At present, the model parameters are estimated for labour market participation in each country by gender and separately for different age groups. This is of key importance for modelling educational participation and attainment since these are known to be gender and age specific. This expanded model framework is then used to create a detailed set of baseline projections for labour supply, disaggregated by country, age groups and gender and covering a 10-15 year period. This model forms a key input for the analysis of the supply of qualifications and provides the link between economic activity and labour market supply. Finally, this link can be used to provide a range of projections of available skills through scenario-based analysis around the baseline forecast, indicating areas that are most sensitive to the economic climate and change.

Modelling and forecasting the supply of qualifications ideally requires a detailed and comprehensive stock-flow model, with behavioural links which can be used to predict the distribution of people in the total population and labour force (employed and unemployed people) by highest qualification held ([4]). In practice, this ideal is hard to realise, as a detailed demographic or educational and labour market accounting system is still lacking at EU level.

The methods currently used for modelling forecasts are less ambitious. They range from rather simple models, based on fitting trends of aggregate

([2]) Skill supply is not measured by occupations as it is not possible to attribute people to different jobs after acquiring particular qualification: the occupational decisions of individuals vary significantly and cannot be predicted. Moreover, employment in occupations will change in the life course of individuals.

([3]) Detailed model description is available at: http://www.camecon.com/AnalysisTraining/suite_economic_models/E3ME.aspx.

([4]) This kind of model is used in many national projections of skill supply, see e.g. Wilson and Bosworth (2006) for the UK.

qualification patterns among the population and/or labour force, to more sophisticated approaches based on econometric analysis of microdata on individuals, mainly using LFS data. All focus on overall stocks rather than flows.

### 1.2.2. Demand for skills

The demand side involves four main elements or modules. Each module contains its own database and models. The results focus on future demand trends at a pan-European level (EU-27+): by sector (up to 41 industries based on NACE classification); by occupation (up to 27 occupations based on ISCO classification); by qualification (three broad levels based on the ISCED classification); plus replacement demands by occupation and qualification. Together these produce estimates of the numbers of job openings (net employment change plus replacement demand) by skill (as measured by occupation and by qualification). The detailed classifications and aggregations used are provided in Annex 2.

The forecast of employment by economic sector is provided by a module which is based on results from the existing pan-European multisectoral macroeconomic model (E3ME). This model delivers a set of consistent sectoral employment projections, which are transparent in terms of the assumptions made about the main external influences on the various countries (including technological change and the impact of global competition).

E3ME combines the features of an annual short- and medium-term sectoral model, estimated by formal econometric methods, with the detail and some of the methods of the computable general equilibrium models that provide analysis of the movement of the long-term outcomes. It can also be used for dynamic policy simulation and for forecasting and projecting over the medium and long term.

The LFS conducted in all countries provides a source of information for the construction of occupation-industry matrices of employment. These surveys have the advantage of being conducted regularly. They also adopt standardised sets of questions and systems of classification. While there are still some differences among countries, LFS provide a broadly consistent set of data which can be used for producing occupational employment projections within the industries identified in macroeconomic models such as E3ME. The forecasting module designed to calculate changes in employment (expansion demand) by occupation (EDMOD) based on these data works out the implications for occupational employment.

Occupational employment patterns are only one way of measuring skills. An occupational category can be understood as broadly describing a particular job (related tasks, requirements, position, etc.). Qualifications represent the characteristics of people filling these jobs as well as one of the selection criteria for filling a particular job. From the education and training policy and planning point of view, the types of qualifications typically required are important. Even with only weak data for (formal) qualifications, it has been possible to develop the module (QMOD) which allows inferences to be made about implications for qualifications.

In addition to changes in overall occupational employment levels, it is important to consider replacement demand arising from outflows from a job/occupation, such as retirements and deaths, transition to non-employment, net migration and inter-occupational mobility. Estimating replacement demand is not straightforward and is quite sensitive to the data sources used. Ideally, detailed data on labour market outflows and transitions (mainly retirements and occupational mobility) would be required to analyse replacement demand more accurately. However, these are not currently available and therefore this forecast relies on a methodology that is based on stocks of age-cohorts by occupation and qualification, and excludes transitions from one occupation to another.

From the LFS, it is possible to analyse the demographic composition of each occupation. This allows specific rates of retirement to be estimated for each occupational class (but still not taking account of inter-occupational mobility). LFS data can also be used to estimate rates of outflow. The replacement demand model (RDMOD) has been developed on the basis of data sources that are similar to the occupational model (EDMOD). The model is driven in part by the occupational and qualification employment levels projected from EDMOD and QMOD, combined with models and information on the probability of leaving employment owing to retirement or migration and for other reasons (e.g. transition to inactivity).

### 1.2.3. Comparing skill supply and demand

To provide information on possible labour market imbalances and skill mismatches, a further module (BALMOD) has been added. This module compares the skill demand and skill supply projections (focusing on qualifications) and attempts to reconcile the two.

The possibility to analyse potential skill imbalances in the labour market is important from a policy and individual point of view. Such information can, in conjunction with corresponding demand estimates, shed light on possible future developments in European labour markets, highlighting potential mismatches and thus helping to inform decisions on investments in skills (especially in formal qualifications) made by individuals, organisations and policy-makers.

However, simply comparing current demand and supply projections is problematic for both practical and theoretical reasons. Although the two sets of results are based on common data and are carried out simultaneously, they do not incorporate direct interactions between supply and demand and, therefore, they cannot be directly compared. Cedefop has started to work on modelling interactions between supply and demand, but due to the complexity of the task these interactions might be incorporated only in the medium to long term. There are various other conceptual and methodological issues regarding imbalances that need to be considered to avoid misleading inferences and interpretations.

A final adjustment has been made to the estimates of employment by qualification (demand side) to take account of the labour market accounts residual. This residual measures the difference between employment as measured for the NA estimates (workplace based, jobs) and the corresponding LFS estimates (heads, residence based). Both measures are used in the project ([5]). The difference between the NA and LFS can be quite significant and needs to be considered, especially when comparing demand and supply.

Differences between skill demand and supply can include:

(a) double jobbing (some people have more than one job) or one full-time job is shared by two or more people;
(b) distinction between residence and workplace (many people do not live in the same country as they work; this is especially significant for some small countries such as Luxembourg);
(c) participants in training and related schemes who are also working in parallel (they may be included in the labour force and in education statistics – double counting);
(d) different definitions of unemployment (e.g. ILO definition versus limited to unemployment beneficiaries);
(e) statistical errors (in measures of employment, unemployment and related indicators, including sampling and measurements errors);
(f) other differences due to the use of different data sources such as treatment of nationals working abroad.

## 1.3. Structure of the publication

This publication presents a compilation of background technical reports which were submitted to Cedefop as part of the research project *Skills needs and supply in Europe*. The main aim is to provide an overview of the methodology

([5]) For a detailed discussion see Kriechel and Wilson (2010) and Livanos and Wilson (2007a).

rather than to present the results. Therefore results are not provided systematically and any results presented are only illustrative and may differ from the actual forecasting results. More detailed information can be found in the list of references or by visiting the Skillsnet members' area on the Cedefop web portal ([6]).

The Chapter 2 of this publication presents the construction of the database used in the Cedefop model. High quality data that are consistent over time and comparable across countries are a prerequisite for such a modelling exercise and for obtaining reliable and relevant results. Chapter 3 provides a detailed overview of the key features and extensions of the underlying macroeconomic model (E3ME). Chapter 4 describes the key elements of the labour supply models and attempts to introduce more of the stock-flow elements into the model. Labour demand is considered in two dimensions in Cedefop's methodology. The modelling of net changes in employment, expansion demand, is described in Chapter 5. The demand for labour created by the outflow of workers from their occupations for various reasons, replacement demand, is the scope of Chapter 6. Even though the skills supply and skills demand modules are developed under the same general framework, the comparison between these two sides of the labour market is not straightforward. Chapter 7 is hence dedicated to describing how to reconcile supply and demand as well as the imbalance indicators developed in the project. The final two chapters of the publication focus on evaluation of the previous pilot forecasts. Chapter 8 focuses on a quantitative assessment of the accuracy of the forecasts. Chapter 9 is devoted to the quality of the forecast based on stakeholders' satisfaction. Chapter 10 summarises all the main findings of each section and outlines the next steps.

([6]) http://www.cedefop.europa.eu/EN/about-cedefop/projects/forecasting-skill-demand-and-supply/membership.aspx [accessed 7.5.2012].

CHAPTER 2.
# Historical data foundation

A coherent and consistent database containing historical time series is a prerequisite for forecasting skill supply and demand. Without data that are consistent over time and reasonably comparable across countries, the forecasts generated will have a little practical relevance, irrespective of the sophistication of the model and plausibility of the assumptions adopted.

At the start of skills supply and demand forecasting at European level, a consistent database was developed comprising the key statistics used as a sound basis for the development of the model. Current work is focused on dealing with various data issues and exploring further detailed breakdowns by category of occupation and education.

The main issues with labour market data concern sample size, missing observations and small numbers in particular cells, for both the demand and supply side of the labour market.. Another important issue concerns the (regular) revisions and updates of international classifications during the development of the methodology. We appreciate the valuable support from Eurostat, which provided us with additional information necessary to successfully deal with such issues.

This chapter is based on working papers by Stehrer and Ward (2010 and 2011). It describes the key adjustments made to construct the improved general database. A key question is which methodological data set is the best to use. The labour force survey (LFS) provides more detailed data that are better suited to of Cedefop’s needs, while national accounts (NA) data are more suitable for general economic modelling. This issue is tackled in Section 2.1. The development of a consistent database for cross-country analysis is discussed in Section 2.2. The adjustments of demand and supply data are the focus of Section 2.3. The cross-country heterogeneity in education and even classification systems makes the production of a harmonised database even more complicated. This is illustrated in Sections 2.4 and 2.5 which analyse the differences by level of education and field of study in the most common occupations across countries.

## 2.1. Labour force surveys versus national accounts estimates: strengths and weaknesses

The LFSs conducted throughout the EU provide an invaluable source of information on employment broken down by industry and occupation. They have the advantage of being conducted much more frequently than a typical census. They also adopt much more standardised sets of questions and systems of classification. While there are still some differences among countries, this source provides a broadly consistent set of data across Europe which can be used to produce occupational employment projections within the industries identified in macroeconomic models.

However, as described in more detail in Livanos and Wilson (2007a), the LFS is not without its faults. In many respects estimates of employment based on NA are more robust and more compatible with other economic indicators. Estimates of employment in NA may differ from the results found from other statistics and surveys, in particular the LFS. These differences arise mainly due to the integration of other sources and for conceptual reasons.

More specifically, NA integrate information from many sources. All available sources (including LFS) are assessed before deciding on the best way of integrating them. Each source may shed light on a different part of the economy. Some countries make little use of the LFS in their NA. The information is combined to provide the most complete and consistent estimate. As a consequence, each individual basic source may provide results that are different from the integrated NA estimates. In NA, employment figures must be consistent with other variables such as output and compensation of employees (i.e. wages, salaries and social contributions).

The European system of accounts (ESA95) acknowledges two employment concepts that depend on geographical coverage: resident persons in employment (i.e. the national concept of employment) and employment in resident production units, irrespective of the place of residence of the employed person (i.e. domestic concept). The difference between them corresponds mainly to the net number of cross-border workers. The domestic concept is more appropriate when examining employment and gross domestic product together. The LFS, on the other hand, covers resident households. Hence, the LFS gives information on a major subset of the national concept. This means that:

(a) LFS data must be adjusted, mainly for cross-border workers, to be consistent with the domestic concept normally used in NA;
(b) appropriate adjustments are needed on coverage: the LFS does not cover persons living in institutional or collective households (e.g. conscripts),

unpaid apprentices and trainees and/or persons on extended parental leave, who are all covered by ESA95 employment;

(c) adjustments must also be made to record thresholds: LFS results exclude persons below 15 years old from the definition of employment (in some countries the exclusion boundaries are below 16 years old and/or above 75 years old). NA do not exclude individuals from employment because of age. The difference is small in developed economies.

The size of these conceptual adjustments is modest, with the possible exception of conscripts, and cross-border workers for small countries.

The LFS data from Eurostat are sample data. For many countries the time series are short and the number of respondents within a particular cell is often low. This means that the estimates of occupational structure within sectors are not always precise or robust. These problems are even more serious when it comes to making estimates of replacement demand, which asks even more of the data. Nevertheless, LFS data are often the only data available for breaking down employment by sub-group of the population. Moreover it has become well established in the analytical and political discourse (e.g. European benchmarks and indicators), so that survey results are becoming more familiar to the general public than those based on the NA.

## 2.2. Developing a consistent database for cross-country comparisons

There are two major issues with the LFS series as a source of detailed employment and labour supply characteristics. These issues must be addressed for the database to be consistent and comparable across countries.

The first issue is the fact that the data represent a sample, and the people surveyed are relatively few in number and their characteristics change from one year to the next. In other words, the survey data are only approximately representative of the population as a whole at certain levels of detail (sector, occupation, education). Moreover, the degree of 'sample variability' varies among countries, which means that at the very detailed level which is required for detailed analysis, variations in the employment recorded in the smaller occupations and sectors are inevitable, especially in less populated countries where the relative smallness of the sample is most problematic.

The second issue is that there are changes in the method of collecting data, especially in the sampling frame but also in the systems of classification used and the way data are collected on the basis of a national classification system which is mapped onto the EU system. The latter applies especially to the data on occupations, where many countries have their own system that they continue to use.

The problem with sample nature is difficult to overcome because the discontinuities in the data series which result are an inherent feature of the LFS data and are, therefore, undocumented. To deal with changes in methodology is in principle easier because in many cases they are documented. However, this does not make it any easier to find viable solutions to building a consistent series. Arriving at a consistent database that allows for easy cross-country comparisons of detailed employment structures, or at least one which is disturbed by a minimum amount of statistical noise, is a challenge.

Another source of discontinuity in data series is even more difficult to overcome: changes in the classification of jobs due to changes in the nature of jobs over time. Many jobs involve a mix of tasks that changes because technological advances and changes in the organisation of work alter the nature of the production process. Consequently, a job that required a certain combination of skills in the past may require a different combination now. In some cases the difference is so great that it can alter the classification of a job within the ISCO system ([7]). These types of change tend to happen gradually over time and come to light only during periodic reviews of the classification system. In some cases, the changes are universal and affect all countries, as in the case of the revision of ISCO (from ISCO 88 to ISCO 08), which involved mainly shifting a few occupations between categories because of the change in skill-set requirements. In other cases, the changes are specific to a particular country and reflect the difficulties of classifying many jobs that could be included in different parts of ISCO. In these cases, there tends not to be a uniform classification and, accordingly, certain occupations are liable to be shifted from one category to another in some countries as particular aspects of the job become more or less important. Whether they affect one country or many, the common characteristic of such changes is that the jobs in question, and the people performing them, shift suddenly from one category to another, or more typically, from one category to several others, making adjustments problematic.

The difficulty is compounded by the fact that the same type of change does not necessarily occur at the same time in all countries because of differences

---

([7]) For example ISCO 08 introduced new category of occupation 3122, Manufacturing supervisors, within broad Category 3 – Technicians and associate professionals. The workers classified within this group were in ISCO 88 categorised in category 8 – Plant and machine operators and assemblers with no difference.

among countries both in the pace of technological progress and in the nature of jobs. Furthermore, since countries are positioned differently in the international division of labour (or in the activities in which they specialise), not only does the mix of tasks involved in a particular sector differ among countries, but the tasks associated with a particular job within a sector also differ. Moreover, areas of specialisation change over time, though not necessarily at the same pace, or even in the same direction, in all countries. To complicate matters further, changes in occupational classification systems are not harmonised across Europe, except when Eurostat decides to initiate a common change in ISCO. Since many countries have their own national systems of classification, they can change at any time, and the changes are then reflected in the mapping to ISCO.

ISCO can be applied for practical purposes uniformly in all countries, irrespective of the underlying cause of the discontinuity. Given the sheer volume of the data which need to be harmonised, such a general approach is the only viable one. Moreover, even if one wanted to distinguish the different causes of discontinuities to be able to apply appropriate methods of adjustment, the information is not available to do so.

## 2.3. Approach to adjusting the data

The database to be constructed for the modelling must meet exacting demands. It requires a level of detail which cannot be readily provided in a coherent way. The basic data received from Eurostat must therefore be adjusted considerably. This is difficult owing to the number of indicators to be combined. The question of the data quality is not confined to one or even two time series. Rather, the data assembled have to be made comparable both across countries and over time and to be internally consistent. The detailed breakdown of occupational and educational categories for each country also created additional work.

The data for different years can be combined in this way with the LFS since generally different people are surveyed each year. There can be some overlap because, to increase the consistency of the quarterly data, a small proportion of the sample surveyed remains the same between quarters.

The first step in the process of data adjustment was to develop a consistent raw data set that accounted for the dimensions to be included in the demand- and supply-side models. This allowed flexible use of the data at later stages of the project (including adjustments to the data). As emphasised below, there is no one way of dealing with methodological breaks over time and methodological differences among countries in the series to guarantee consistency. Though the time dimension is less relevant as the approach focuses only on a few, more recent years (during which fewer breaks occurred), there is still the problem of having inconsistent and incomparable detailed data across countries.

To obtain data from the real microdata of the EU-LFS ([8]) in such detail required Eurostat statisticians to use a particular extraction procedure covering the dimensions needed. After a negotiation process leading to agreement to the extraction, the data obtained had to be processed and analysed in an appropriate way. Therefore, historical data were first collected for each year from 1995 to 2008 ([9]).

The data extracted covered:

(a) the demand for labour by gender, broad age group and educational attainment level (based on ISCED-97). It could also be divided by ISCO88 3-digit occupational group and by NACE, Rev 1, 2-digit sector of activity together with the field of study;

(b) the potential supply of labour, defined as the population of working age (i.e. those aged 15 and over), by employment status, by broad educational attainment level (based on ISCED-97) and by ISCO 3-digit (for those in employment) for each year for those aged 15-69 and for those aged 70+.

### 2.3.1. Labour demand data

The necessary historical data were extracted from the EU-LFS and came from the full microdata set maintained by Eurostat. These cover EU-27+. With regard to employment (i.e. the data to be used as the basis for projecting the demand for labour), the data relate to each year between 1995 and 2008 for which LFS data are available, and cover (not including missing categories):

(a) 60 sectors (NACE Rev 1 2-digit);

(b) 149 occupations (ISCO 3-digit)

(c) 7 educational categories (ISCED categories 0-1, 2, 3 short, 3 other, 4, 5, and 6);

(d) age groups (15-54, 55-64, 65-69, 70+);

(e) gender variables.

In addition, the variable ‘field of study’ includes 15 categories listed in Table A4 of Annex 2.

([8]) The microdata delivered by Member States contained all of the detail included in the survey, unlike the published version in which many dimensions are included only in aggregated form.

([9]) Annual averages were collected from 2002 onwards when the quarterly series began, since before then the data were collected in the second quarter of each year.

Potentially, therefore, the data consist of 500 640 cells per country per year (not considering the variable 'Field of study' which is treated separately) ([10]). For efficient further use, the original data set was split into one file for each country, together with a harmonisation of codes and variable names. The data were also organised such that the resulting country files included the whole set of potential cells with respect to the characteristics described above (i.e. 500 640 cells per year per country). This enables a better comparison to be made across countries and over time and facilitates further processing. For the analysis performed in 2011, the data have been restricted to include years starting from 2004 onwards.

The next step was to consider missing information with respect to some categories (e.g. educational attainment and occupation) because of non-response. These observations are referred to as 'randomly missing' information. In such cases, the number of persons with 'no response' were allocated pro rata on the basis of the distribution of the missing information between the other dimensions. For example, when the information on educational attainment was not available for a particular employed person, it was imputed according to the educational structure in other dimensions, such as industry (NACE 2-digit), occupation, age group and gender. Other problems arose for earlier years. For example, for occupational categories for some countries and for some years (especially earlier ones) only the 1-digit information level was provided and information on some dimensions was completely missing all. However, these problems are no longer present in the data for 2004-08 ([11]).

There are several further data problems that needed to be tackled. First, the small sample size of the EU-LFS (Table 1) raises issues of data reliability, in particular for cells with low numbers of employed people. This becomes more problematic as one considers further detailed breakdowns.

Secondly, there are documented breaks in the series raising the issue of comparability over time and, apart from these documented breaks, there are also several other breaks which become visible only when examining the data in some detail. This was documented in details in Stehrer and Ward (2009).

---

([10]) In practice the number is even larger if the 'no response' category and additional NACE categories for some countries (e.g. NACE-97 and NACE-99) are included.

([11]) The data are made available to other participants in the study in the form of Excel files with dimensions broken down by NACE 2-digit, occupation, education, gender and age.

Table 1. **Sample size of EU-LFS, 2007**

| | Average number per quarter | % of population 15-74 | | Average number per quarter | % of population 15-74 |
|---|---|---|---|---|---|
| BE | 22 100 | 0.3 | LU | 4 000 | 1.1 |
| BG | 26 000 | 0.4 | HU | 60 100 | 0.8 |
| CZ | 49 300 | 0.6 | MT | 5 600 | 1.8 |
| DK | 20 700 | 0.5 | NL | 83 400 | 0.7 |
| DE | 131 300 | 0.2 | AT | 37 900 | 0.6 |
| EE | 4 800 | 0.5 | PL | 40 000 | 0.1 |
| IE | 59 100 | 1.8 | PT | 33 100 | 0.4 |
| EL | 56 100 | 0.7 | RO | 49 900 | 0.3 |
| ES | 126 300 | 0.4 | SI | 13 900 | 0.9 |
| FR | 63 900 | 0.1 | SK | 21 900 | 0.5 |
| IT | 127 200 | 0.3 | FI | 34 400 | 0.9 |
| CY | 7 300 | 1.3 | SE | 51 300 | 0.8 |
| LV | 7 500 | 0.4 | UK | 86 900 | 0.2 |
| LT | 12 400 | 0.5 | CH | 43 400 | 0.8 |

*Source*: EU-LFS.

Thirdly, there is a problem of missing numbers for particular observations. In many cases a cell (meaning a cross combination of NACE industry, education, occupation, age and gender) is reported for only one year over the whole period considered.

There is no easy solution to the discontinuities since there are no data present to provide a link. Only in very few cases it is possible to detect shifts offsetting a particularly large change in one dimension – e.g. a large increase in some ISCO 2-digit occupation group might coincide with a large decrease in another. However, such offsetting shifts mainly tend to be spread across sectors or occupations and in some cases sectors, occupations and educational attainment levels are all affected. An adjustment for pairwise shifts might therefore work sometimes but cannot be a general strategy.

For occupation data, the incidence of discontinuities, and consequently the need for adjustment, was particularly high, reflecting the difficulty of determining the appropriate ISCO category in which to classify particular jobs ([12]). Moreover, the nature of jobs tends to change over time, so adding to the chances of discontinuities in the series. As in the case of other dimensions, the extent of discontinuities in the occupation series varies markedly among countries. In some, it is a relatively minor problem while in others it is a major flaw affecting a significant share of total employment. However, most or all of the breaks occurred prior to 2004 (the year in which this analysis starts), so this problem is less relevant. However, a relatively wide scope still exists for differences in classification among countries, which is hard to deal with.

### 2.3.2. Labour supply data

The historical data for this task have been extracted from the same source as the data for demand, and they cover the same countries for the same years. The data cover the following dimensions:

(a) employment status (employed, unemployed, inactive),
(b) 149 occupations (ISCO 3-digit) for the employment status 'employed',
(c) seven educational categories (ISCED categories 0-1, 2, 3 short, 3 other, 4, 5, and 6),
(d) 56 age groups (55 individual years of age 15-69 and 70+),
(e) gender variables.

The age groups have been combined into five-year age cohorts 15-19, 20-24, ..., 65-69, 70+) resulting in 12 groups. This gives 25 032 potential cells per country per year for those employed and 168 potential cells for those unemployed and inactive. Data have been processed and adjustments made for non-responses.

It is important to consider the provision of more details for educational and occupational categories (from ISCO 2-digit to ISCO 3-digit). To get reliable estimates of the shares of the combinations considered ([13]), information only for the years 2004-08 was used. These data seem to be more reliable as the methodological breaks are limited. The average of employed persons over these years was calculated for each combination. A question arises whether missing values in particular years for observations with positive values in other years should be treated as missing or as zero values.

---

([12]) For labour demand the combination used was NACE-ISCO-EDU (4 categories)-age-gender.

([13]) For labour supply the combinations were ISCO-EDU-age-gender and EDU-age-gender.

Another strategy would give less weight to irregularly observed cells (as means are calculated including the zero values) though maintaining the overall average; it would give more weight to those irregularly observed cells but would result in an increase in the average overall number of persons employed. Furthermore, it would be more consistent with the approach taken for the adjustments over time. Calculations are done for both cases; the results reported below are based on the first strategy (differences have still to be checked in detail).

## 2.4. Variations in the education level of occupations among countries

The current information on the qualification level for both the demand and supply sides of the labour market is restricted to three broad levels of formal qualification. There is justified demand from the stakeholders for more detailed information on educational levels and fields of study for all occupations. As previously mentioned, the key precondition of such an exercise is to have appropriate data which are consistent for all countries and over time. As we are facing the similar problems as those mentioned above, an initial analysis to assess the feasibility of finding some general pattern could help to have the data in the required form.

The nature of the variations in the education level of those performing particular jobs is shown by considering the differences among specific countries and how far they are related to the supply of people of working age with particular educational qualifications. Five ISCO 3-digit occupations in which a significant number of people are employed in all countries and six Member States with different features and different levels of economic development are selected here for illustrative purposes. The five occupations are manager of a small business (ISCO 131), business professional (ISCO 241), administrative associate professional (ISCO 343), building frame and related trades worker (ISCO 712) and motor vehicle driver (ISCO 832). The six countries are Bulgaria, the Czech Republic, Germany, Spain, Italy and the UK.

First, however, it is instructive to compare the numbers of people of working age (15-64) with different levels of educational attainment, defining this in terms of the main ISCED categories, which gives a general indication of the potential workforce from which employers can draw ([14]).

By taking the 15-64 age group, we include people who are still in the process of completing their education and initial vocational training even though they may be in work. Their inclusion may give a misleading impression of the education levels required to perform a particular job because some of the workers included will be recorded, for example, as having only basic schooling whereas they are in the process of completing an upper secondary level training programme which they need to do to acquire the necessary skills and competences. The same may apply to some of those with upper secondary education who are taking a tertiary-level programme. This may again be required as a means of acquiring the skills needed for the job.

The implication is that there is a need to be able to distinguish young people in the employment data set who might be in this position, though at present it is not possible to do so (data set at present the focus is on distinguishing the older members of the workforce in employment). A more disaggregated data set will, therefore, be requested from Eurostat so that this aspect can be considered in future analyses.

#### 2.4.1. Working age population by detailed level of education

The potential workforce differs markedly across the six countries in terms of the numbers with different education levels. In particular, the proportion of those with no qualifications beyond basic schooling (ISCED 1, 2 and 3c) varies from almost 50% in Italy to around 30% in the UK and Bulgaria to under 10% in the Czech Republic. Equally, the proportion with tertiary qualifications varies from around 27-29% in Spain and the UK (countries in which the number of people with only basic schooling is also relatively large, indicating a strongly diversified workforce) to only 13% in the Czech Republic and 11% in Italy (countries with very different numbers of low-educated individuals).

([14]) It is arguable that a wider definition of the workforce than the one conventionally used, namely the employed plus the unemployed, is more meaningful, since a significant number of those labelled as inactive tend to be ready to take up employment if work is available. This explains why those previously recorded as inactive and unemployed take up the additional jobs created when employment grows. It also explains why the number of self-reported unemployed individuals tends to be markedly larger in some cases than the number recorded as unemployed according to the standard ILO definition – i.e. available for work and actively seeking work.

Table 2. **Working age population (15-64) by educational attainment**

*(%)*

| | BG | CZ | DE | ES | IT | UK | EU-27 |
|---|---|---|---|---|---|---|---|
| 1 | 5.4 | 0.8 | 5.9 | 19.5 | 14.2 | 0.3 | 9.5 |
| 2 | 24.8 | 9.1 | 17.5 | 26.2 | 34.2 | 14.8 | 21.3 |
| 3c | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 13.8 | 1.7 |
| 3o | 51.4 | 75.0 | 49.4 | 24.4 | 39.0 | 43.7 | 44.2 |
| 4 | 0.6 | 1.8 | 6.3 | 0.2 | 1.1 | 0.1 | 2.7 |
| 5 | 17.5 | 12.8 | 19.7 | 28.5 | 10.8 | 26.4 | 19.8 |
| 6 | 0.3 | 0.5 | 1.2 | 0.7 | 0.1 | 0.9 | 0.7 |

*Source*: Stehrer and Ward (2011).

Accordingly, the proportion of the working age population with upper secondary education, which in practice means mainly those who have successfully completed a training programme of at least three years for a particular vocation, is particularly large in the Czech Republic (almost 80%) and especially small in Spain (only around 25%).

#### 2.4.2. Detailed education levels within selected occupations

Small business managers (ISCO 131) have on average a relatively high level of education. The proportion with tertiary education is above the EU average in Spain and the UK (23-24%), where a relatively large number of people of working age have this level of education. However, it is much higher than average in Bulgaria (39%) where the share of the working age population with tertiary education is below average. It is also higher than average in the Czech Republic (28%) where the share of the working age population with this education level is even smaller. Moreover, in both countries almost all of those employed in this occupation who do not have tertiary education have upper secondary qualifications. This contrasts with Italy and Spain, where around half only have basic schooling. Italy stands out with only a small proportion (6%) of those employed as small business managers having tertiary education, which is in line with the small proportion of the working age population with such qualifications.

Indeed, there is clearly a relatively strong correlation in all these six countries between the educational breakdown of those employed as small business managers and that of the working age population. Much the same applies in all EU Member States.

Table 3. **Educational levels by occupation**

| ISCO | BG | CZ | DE | ES | IT | UK | EU-27 |
|---|---|---|---|---|---|---|---|
| **Small business manager (131)** | | | | | | | |
| 1 | 0.2 | 0.0 | 2.0 | 20.6 | 8.6 | 0.2 | 10.0 |
| 2 | 3.0 | 1.5 | 12.8 | 31.3 | 37.5 | 12.3 | 20.3 |
| 3c | 0.0 | 0.0 | 0.0 | 0.2 | 0.3 | 14.0 | 2.0 |
| 3o | 56.7 | 68.1 | 58.1 | 25.2 | 47.1 | 49.1 | 45.0 |
| 4 | 0.7 | 1.6 | 8.5 | 0.1 | 0.7 | 0.1 | 2.7 |
| 5 | 39.1 | 28.4 | 18.2 | 22.5 | 5.7 | 23.8 | 19.8 |
| 6 | 0.3 | 0.4 | 0.5 | 0.1 | 0.0 | 0.5 | 0.3 |
| **Business professional (141)** | | | | | | | |
| 1 | 0.0 | 0.0 | 0.3 | 0.0 | 0.4 | 0.3 | 0.3 |
| 2 | 0.1 | 0.3 | 3.4 | 0.0 | 3.4 | 0.8 | 2.0 |
| 3c | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 4.4 | 0.8 |
| 3o | 13.4 | 50.9 | 24.7 | 1.1 | 41.6 | 28.3 | 23.1 |
| 4 | 0.2 | 2.5 | 10.4 | 0.0 | 1.0 | 0.1 | 3.8 |
| 5 | 86.3 | 45.9 | 59.1 | 96.9 | 53.3 | 65.8 | 68.6 |
| c6 | 0.0 | 0.4 | 3.2 | 1.8 | 0.3 | 1.1 | 1.4 |
| **Administrative associate professional (343)** | | | | | | | |
| 1 | 0.0 | 0.0 | 0.1 | 1.4 | 0.3 | 1.1 | 0.5 |
| 2 | 0.5 | 0.6 | 6.6 | 8.6 | 7.4 | 1.1 | 6.2 |
| 3c | 0.0 | 0.0 | 0.0 | 0.3 | 0.2 | 8.5 | 0.3 |
| 3o | 65.1 | 81.3 | 60.0 | 33.3 | 75.7 | 39.8 | 55.6 |
| 4 | 2.0 | 3.1 | 15.4 | 0.1 | 2.5 | 0.5 | 6.8 |
| 5 | 32.2 | 14.8 | 19.4 | 56.0 | 13.7 | 48.2 | 30.4 |
| 6 | 0.1 | 0.2 | 0.3 | 0.2 | 0.0 | 0.8 | 0.2 |
| **Building frame and related trades worker (712)** | | | | | | | |
| 1 | 2.6 | 0.0 | 2.7 | 31.0 | 24.0 | 0.1 | 17.7 |
| 2 | 33.5 | 5.0 | 13.9 | 44.8 | 56.1 | 12.7 | 27.7 |
| 3c | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 11.0 | 1.4 |
| 3o | 62.0 | 94.6 | 68.8 | 16.5 | 18.8 | 69.4 | 47.0 |
| 4 | 0.6 | 0.1 | 1.8 | 0.1 | 0.1 | 0.0 | 1.3 |
| 5 | 1.4 | 0.2 | 11.3 | 7.3 | 0.8 | 6.6 | 4.7 |
| 6 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.1 | 0.0 |
| **Motor vehicle driver (832)** | | | | | | | |
| 1 | 0.2 | 0.0 | 1.9 | 22.2 | 12.6 | 0.2 | 9.7 |
| 2 | 14.4 | 6.6 | 17.9 | 46.0 | 58.9 | 18.7 | 26.6 |
| 3c | 0.0 | 0.0 | 0.0 | 0.2 | 0.5 | 18.6 | 2.7 |
| 3o | 81.9 | 92.6 | 73.9 | 23.5 | 27.0 | 57.4 | 55.6 |
| 4 | 0.3 | 0.2 | 2.6 | 0.1 | 0.4 | 0.1 | 1.5 |
| 5 | 3.3 | 0.6 | 5.2 | 8.0 | 0.7 | 4.7 | 3.9 |
| 6 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 |

*Source*: Stehrer and Ward (2011).

Business professionals (ISCO 241) have a much higher level of education than small business managers. As with the small business ma, however, there is some variation among the six countries in the proportion of employed business professionals with a tertiary education, which is not entirely in line with the variation in the educational breakdown of the working age population. The proportion with tertiary qualifications is virtually 100% in Spain, which is much higher than in the UK (67%), a country with a similar share of working age population with this level of education. It is also very high in Bulgaria (86%), which again is out of line with the relatively small share of highly educated members of the potential workforce. On the other hand, the proportion with tertiary education is relatively small in the Czech Republic and Italy. In all countries, a very small proportion have only basic schooling, so that virtually all of those employed as business professionals have at least upper secondary qualifications. The big difference among countries is in the split between those with upper secondary qualifications and those with tertiary level.

A smaller proportion of administrative associate professionals (ISCO 343) have tertiary level education in all countries and a slightly larger proportion has only basic schooling. In this case, therefore, most of those employed in the EU have upper secondary qualifications. This is the case in four of the six countries, the exceptions being Spain and the UK, both of which have a relatively large share with tertiary education and in both of which the share of the working age population with this level of education is also large. Indeed, for this occupation, the division of employment by education level is more in line with the division of the working age population, with both the Czech Republic and Italy having a relatively small share with tertiary education, though this is also the case for Germany. In Bulgaria again the proportion of people with tertiary qualifications employed in this occupation is higher than the EU average and a much higher than in Germany, despite the fact that Bulgaria has a smaller proportion of the potential workforce with such qualifications.

Most building frame and related trades workers (712) also have upper secondary level qualifications in most countries, but relatively few have tertiary education. The exceptions are Spain and Italy, where most workers (76-80%) have no qualifications beyond basic schooling, in line with the relatively large share of the working age population with low education. In Germany, a relatively large proportion of workers have tertiary education, even though it is almost certainly not a requirement for carrying out the tasks involved in the job. On the other hand, the proportion with tertiary education in this occupation in Bulgaria and the Czech Republic is particularly small (almost zero in the Czech Republic and only just over 1% in Bulgaria), in contrast to the situation in the higher level occupations, especially in Bulgaria. As for managers of small businesses, there is a strong correlation between the educational breakdown of those employed in this occupation and the educational composition of working age population.

As building frame workers, most motor vehicle drivers in the EU have upper secondary education, except in Spain and Italy, where around 70% of those employed in this occupation have only basic schooling. This is a lower – albeit still substantial – proportion than for building frame workers, and is in line with the large number of people of working age with a low level of education. Indeed, the variation among the six countries in the breakdown by education for this job is very much in line with the breakdown by education for the working age population. However, in both Bulgaria and the Czech Republic a smaller proportion of those employed in this occupation have only basic schooling than is the case for the working age population. This reflects the disproportionate numbers of people with this level of education who are not in employment at all in these two countries.

The relative number of people of working age with different levels of education constitutes the potential workforce that employers have to recruit from. Therefore, this share plays an important part in influencing the division of employment among workers with different educational qualifications in particular jobs. This, however, is much more the case for some jobs than for others. Nevertheless, especially for some jobs, this is far from being the only factor at work. The question then is to what extent are these variations due to differences in the nature of the jobs rather than to differences among countries in the educational profile expected of workers to do a particular job and in the training programmes they are expected to complete.

## 2.5. Variations in the field of study in selected occupations

The second analysis focused on examining the differences in fields of study in selected occupations. It should also provide deeper insight into the educational profile required for particular jobs by examining the fields of study of employed people during their upper secondary or tertiary education. However, as for education levels, the division of the potential workforce, as defined above, between broad fields of study, should be considered.

Based on the results in Table 4, the most common field of study of those with upper secondary and tertiary education is engineering, manufacturing and construction, which was accounted for around 20% of all those of working age in the EU. This was followed by social sciences, business and law (16%), general programmes (7%), health and welfare (6%) and services (5%).

However, there are substantial variations among countries in the relative numbers completing the different programmes. In particular, in both the Czech Republic and Bulgaria many people of working age studied engineering, manufacturing and construction (42% and 31%, respectively), while the number was also well above average in Germany (25%). By contrast, only 8% of the working age population undertook this programme of study in Spain and 11% in Italy. In Germany, a higher than average proportion (21%) studied social science, business and law programmes, which was also the case in the Czech Republic and the UK. In Bulgaria especially, and to a lesser extent in Spain, a much smaller proportion of the working age population has completed these types of programmes. In both of these countries, a larger proportion than elsewhere has completed general programmes.

Table 4. **Working age population by field of study**

*(%)*

| | | BG | CZ | DE | ES | IT | UK | EU-27 |
|---|---|---|---|---|---|---|---|---|
| 1 | Agriculture and veterinary | 2.9 | 4.6 | 2.0 | 0.7 | 1.1 | 1.2 | 2.2 |
| 2 | Computer science | 0.3 | 0.3 | 0.7 | 1.6 | 1.0 | 1.8 | 1.0 |
| 3 | Engineering, manufacturing and construction | 31.4 | 42.0 | 25.0 | 8.2 | 11.1 | 14.9 | 19.6 |
| 4 | Foreign languages | 0.3 | 0.2 | 0.6 | 0.0 | 1.5 | 1.1 | 0.7 |
| 5 | General programmes | 14.7 | 6.2 | 5.4 | 16.3 | 3.7 | 0.6 | 7.4 |
| 6 | Health and welfare | 2.9 | 4.8 | 7.3 | 4.4 | 2.3 | 9.9 | 5.5 |
| 7 | Humanities, languages and arts | 1.5 | 2.1 | 2.3 | 2.8 | 4.2 | 7.1 | 3.4 |
| 8 | Life science (including biology and environmental science) | 0.2 | 0.3 | 0.3 | 0.5 | 0.6 | 1.7 | 0.6 |
| 9 | Mathematics and statistics | 0.2 | 0.2 | 0.2 | 0.3 | 0.3 | 0.8 | 0.4 |
| 10 | Physical science (including physics, chemistry and earth science) | 0.4 | 1.0 | 0.6 | 1.1 | 0.4 | 2.0 | 0.8 |
| 11 | Services | 2.9 | 8.0 | 6.3 | 2.0 | 2.1 | 5.2 | 4.9 |
| 12 | Social sciences, business and law | 8.5 | 17.0 | 21.4 | 11.7 | 15.3 | 18.1 | 16.0 |
| 13 | Teacher training and education science | 3.5 | 3.3 | 4.1 | 3.1 | 3.8 | 4.9 | 3.4 |
| 14 | Computer use | 0.0 | 0.1 | 0.3 | 0.0 | 0.1 | 1.0 | 0.3 |
| 15 | Science, mathematics and computing | 0.0 | 0.0 | 0.0 | 1.0 | 3.6 | 0.9 | 1.2 |
| NA | | 30.2 | 9.9 | 23.5 | 46.2 | 48.9 | 28.9 | 32.6 |

*Source*: Stehrer and Ward (2011).

The other main points which stand out when comparing fields of study across countries are the relatively large numbers in the UK (as compared with the EU average) completing health and welfare programmes and humanities, languages and arts, compared with the relatively small numbers in Bulgaria, and the relatively large number completing science, maths and computing programmes in Italy. Otherwise, the proportion of the working age population that has studied the different programmes is relatively similar in all countries. (A more detailed breakdown by field of study within selected occupations and countries can be found in Annex 3).

As for education levels, the education and training programmes undertaken by people working in particular occupations reflect to a large extent the proportions of people of working age who have completed the different types of programme in the country in question. Because these proportions vary across countries, so do the education and training programmes. Nevertheless, for certain occupations, a particular type of programme predominates, so that most people working in the jobs concerned have completed the programme in the same field of study, irrespective of the country concerned.

CHAPTER 3.
# Modelling general macroeconomic and sectoral trends

The pan-European skills demand and supply forecast is based on the macroeconomic multisectoral and multi-country model E3ME developed by Cambridge Econometrics ([15]). It is designed to provide a consistent forecast of economic trends, resulting in basic labour demand and supply predictions. It is a large-scale econometric model covering all EU Member States, plus Norway and Switzerland (EU-27+) ([16]), and a detailed disaggregation of 42 economic sectors, consistent with the NACE 2-digit classification (Annex 2). Interaction among economic sectors takes place through input-output relationships and links among countries are formed through international trade equations.

The first section describes the general characteristics of the model. The focus on labour market within the framework of E3ME is described in Section 3.2, which is based on the work of Pollitt et al. (2010). Technological progress and skills upgrading are considered to be important factors influencing skills demand in the model (Pollitt and Chewpreecha, 2010b; Dupuy and Sauermann, 2011). Attempts to incorporate these variables into the modelling exercise are described in the last two sections (Sections 3.3 and 3.4).

## 3.1. General characteristics of the model

The structure of the model is based on the ESA95 (Eurostat, 1996) and includes detailed two-way links among European economies, energy systems and the environment. The economic system is closely tied to the model's treatment of Europe labour markets.

The econometric specification of the model makes it suitable for short and medium-term forecasting and policy analysis. E3ME is estimated and solved on an annual basis, with historical databases covering the period starting in 1970 up to the most recent year that data are available.

---

([15]) For more information see http://www.e3me.com. A full online technical manual is available at http://www.camecon-e3memanual.com/cgi-bin/EPW_CGI.

([16]) The construction of the model allows for extension to a range of countries and currently the possibility to include Iceland, FYROM, Croatia and Turkey are piloted.

E3ME uses a hierarchical system of data sources, with preference given to those that use definitions which are consistent with ESA95 and across national boundaries.

The primary source of economic data is the Eurostat NA branch. Even when Eurostat data are incomplete or believed to be of poor quality, the Eurostat definitions are adopted and the data are completed from other sources. The main example of this is the OECD Stan database, which allows additional disaggregation of some sectors. Other data sources include the European Commission's annual macroeconomic (AMECO) database, as well as information available from the World Bank and the International Monetary Fund. All time series data are collected on an annual basis with gaps filled in using custom algorithms developed using the Ox software application (Doornik, 2007). For the EU-15 (plus Norway and Switzerland) the historical database covers the period from 1970; for the 2004 accession countries data cover period from 1993; and for Romania and Bulgaria coverage starts in 1995. The model uses permanently updated databases so the upper demarcation of the time series depends on the vintage actually used. The main cross-section data consist of input-output tables and bilateral trade matrices, which are sourced from the most recent OECD publications (Yamano and Ahmad, 2006).

The parameters of the model are estimated empirically using formal econometric techniques. In particular, the method utilises developments in time-series econometrics, in which dynamic relationships are specified in terms of error-correction models (ECM) that allow dynamic convergence to a long-term outcome. The specific functional form of the equations is based on the econometric techniques of cointegration and error correction, particularly as promoted by Engle and Granger (1987) and Hendry et al. (1984).

A shrinkage technique is used to model the long-term transition of the 12 recent accession countries so as to match behavioural patterns with the rest of Europe. This effectively fixes long-term parameters in these countries to match EU-15 averages, reducing the reliance on relatively short time series. For further details see Spicer and Reade (2005).

## 3.2. Treatment of the labour market

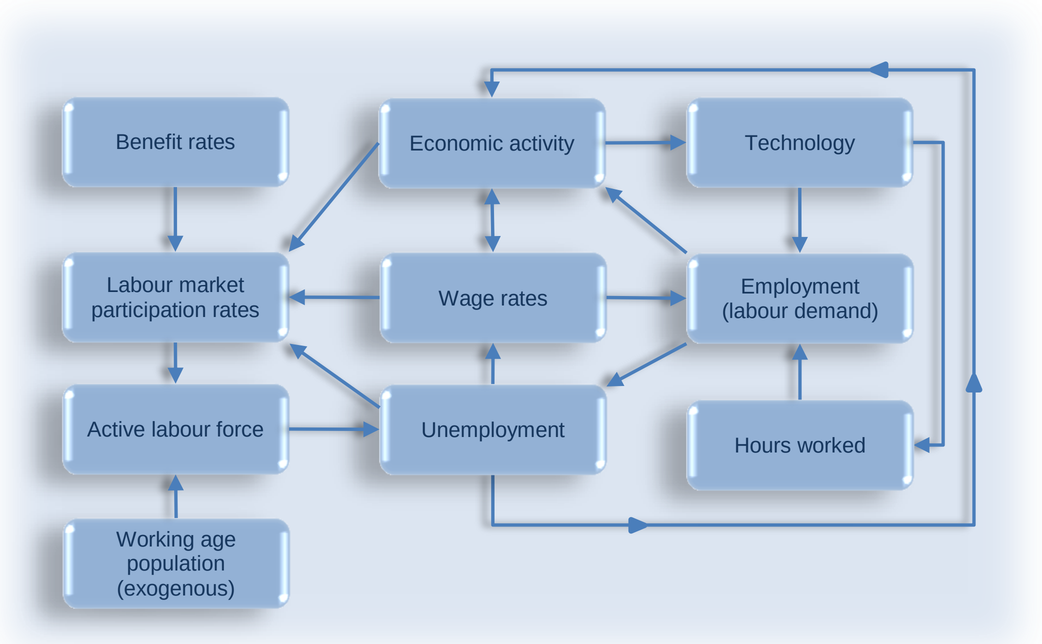
The main role of E3ME in the forecasting project is to describe the links between the labour market and the wider economy. However, because of its detailed sectoral disaggregation, the model is also able to include a relatively complex treatment of the labour market, although it does not directly address skills and qualifications requirements or availability.

The labour market module is made up of four sets of equations: employment demand, average wages, average hours worked and participation rates. Figure 2 shows the key interactions within the module. For a detailed specification of the equations see Pollitt et al. (2010, Annex B).

The original version of E3ME had only a fairly crude treatment of labour supply. This distinguished the supply of labour by gender but not age. This gap was filled by new modelling work conducted by Cedefop as described in detail in Pollitt and Chewpreecha (2008). E3ME has now been augmented to include a detailed treatment of labour supply by age and gender.

E3ME enables the production of a set of benchmark projections of labour supply and demand that provide a coherent European perspective. The use of the macroeconomic model sets the aggregate context for the overall skill demand and supply projections.

Figure 2. **Interactions of labour market in E3ME**




*Source*: Pollitt et al. (2011).

Employment is modelled for each industry and region as a function of industry output, wages, hours worked, technological progress and energy prices. Industrial output is assumed to have a positive effect on employment, while the effect of higher wages and longer working hours is assumed to be negative. The effects of technical progress are ambiguous, as investment may create new jobs or replace labour by automated processes; this varies among sectors.

Hours worked is defined as an average across all workers in an industry, so it incorporates the effects of higher levels of part-time employment in certain regions and industries. Hours worked is a simple equation, where average hours worked by industry and region is a function of 'normal hours worked' (the employees' expected hours based on hours worked in other industries and regions) and technological progress. It is assumed that the effects of technical progress gradually reduce average hours worked over time as processes become more efficient. The resulting estimate of hours worked is an explanatory variable in the employment equation.

In E3ME wages are determined by a complex union bargaining system that includes both worker productivity effects and prices and wage rates in the wider economy. Other important factors include unemployment, tax rates and cyclical economic effects. Generally it is assumed that higher prices and productivity will push up wage rates, but rising unemployment will reduce wage demands. A single average wage is estimated for each region and sector.

The estimates of average wages are a key input to both the employment equations and the price equations in E3ME. If output does not increase, rising wages increase overall unit costs and industry prices. These prices may get passed on to other industries (through the input-output relationships), building up inflationary pressure.

E3ME includes a set of equations for labour market participation rates. When these are multiplied by the exogenous projections for the working age population we obtain a measure of labour supply. The participation rates are affected by economic output, wage rates, hours worked, unemployment, benefit and pension rates, qualifications and the ratio of service activity to manufacturing. As there is little historical precedent, changes in official retirement ages are not directly included in the equations but are taken account of exogenously. At present there is no direct link with official retirement ages. Equations are estimated and solved for each country, disaggregated by gender and by five-year age bands. The separate results for males and females indicate different patterns of activity rates between the two genders, with the activity rate for males being, in general, considerably higher than for females

E3ME does not assume market clearing in the labour market and unemployment arises when labour supply is greater than labour demand. The model database contains time series of unemployment data that are consistent with ILO definitions.

Unfortunately the differences in definitions (between NA and LFS) mean that unemployment cannot be taken simply as the difference between labour supply and demand, and a residual value must be used to make the results consistent with the historical figures. This topic is elaborated more in the Chapter 7.

## 3.3. Involving technological changes into the model

In recent decades technological progress has considerably affected the labour market. Initially, technological progress can be incorporated in the modelling framework through the ICT technological progress indicator which appears in nine stochastic equations and interacts with a skill variable. These indicators are complementary in that new ICT technology requires ICT skills for its operation. However, since the skills variable reflects only ICT skills, it is insufficient for the analysis of the Cedefop forecasting project, which aims to establish links between a wider definition of skills and technologies.

In most age groups a higher level of skills indicates a higher propensity to work. However, for some age groups the reverse is true. For example, in the youngest age groups a higher level of skills naturally obliges people to stay longer in education. Skills are currently treated as exogenous in E3ME. A similar exogenous treatment of technology progress is often found in other types of models, mostly because of its non-measurable nature. There is, nevertheless, scope to endogenise the variable by completing the loop of skill-supply modelling in the model structure within Cedefop's framework.

The approach used to model technological progress in E3ME is adapted from that of Lee et al. (1990). It adopts a direct measure of technological progress by using cumulative gross investment, which is altered using data on R&D expenditure. There are two types of technological progress estimated: ICT and non-ICT.

The equation for the technological progress indicator (ICT, in the equation denoted by *T*) is written as,

$$T_t = c + \alpha \partial_t(\tau 1) \quad (1)$$

where $\partial_t(\tau_1)$ satisfies the following recursive formula

$$\partial_t(\tau_1) = \tau_1 \partial_t - \tau_1 + (1 + \tau_1) \log(GI_t + \tau_2 RD_t) \quad (2)$$

where

$GI_t$ = the level of gross ICT investment;

$RD_t$ = constant-price research and development expenditure;

$\tau_1$ = a measure of the impact of past-quality adjusted investment on the current state of technical advance; while

$\tau_2$ = a measure of the weight attached to the level of R&D expenditure.

To initialise the recursive process for $\partial_t$, the assumption is made that in the pre-data period the process generating $log(GI_t)$ is characterised by a random walk. Under this assumption the first value of *t* can be written as:

$$\partial_0 = log(GI) \qquad (3)$$

where $log(GI)$ represents the average of gross investment over the first five-year sample period. The value of $\tau_1$(the path of technical progress) was set at 0.3 and $\tau_2$ (the relative importance of R&D (*RD*) compared to gross investment (*GI*)) was set at 1.0 (more sophisticated procedures could have been adopted, i.e. a grid-search method based on log-likelihood values). The series $\partial_t(\tau_1)$ is then calculated by working the recursive procedure forward given the initial value 0.

The calculation for the non-ICT technological progress indicator is similar to Equation 2 with two exceptions, namely $GI_t$ is replaced with non-ICT investment and there is no $\tau_2$ term.

Table 5. **Relationship of technological progress indicators with main E3ME variables**

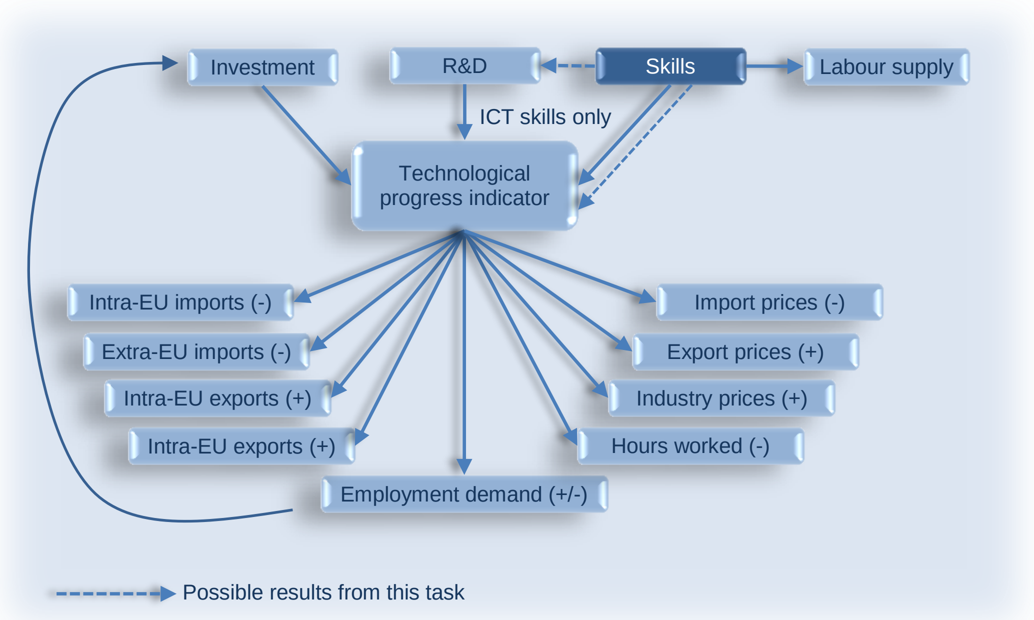
| E3ME's stochastic equation | Relationship with technological progress indicator | Rationale (other things being equal) |
|---|---|---|
| Employment demand | +/- | New technology can be either labour saving or labour creating (e.g. requires skilled worker for its operation) |
| Hours worked | - | New technology increases productivity therefore less hours are required |
| Industry prices | + | Technology increases the quality of goods, allowing firms to charge more for their products |
| Intra-EU imports | - | High technology in the domestic country implies goods are of better quality than elsewhere and therefore there is less demand for imports |
| Extra-EU imports | - | As above |
| Intra-EU exports | + | High technology in export country implies goods are better quality than elsewhere and therefore increases demand for the country's exports |
| Extra-EU exports | + | As above |
| Import prices | - | Imports are of lower quality than high-technology domestically produced goods, implying lower prices |
| Export prices | + | Opposite of the above |

*Source*: Pollitt and Chewpreecha (2010a).

Because E3ME already contains a link between skills and technology, the methodology employed seeks first to improve the current treatment and secondly to achieve a better modelling of skills with technological progress.

The diagram below shows how skills feed into the technological progress indicator in the existing E3ME framework.

Figure 3. **Interactions of skills and technology progress indicator in E3ME framework**




*Source*: Pollitt and Chewpreecha (2010a).

The skills variable (YRKS) that currently appears within the technological progress indicator (YRKC) in E3ME's stochastic equations is initially replaced with a similar skills variable in the E3ME's labour participation rate equation, i.e. a proxy based on the qualification mix (YRKN).

The following function specifications are hence estimated with the new definition of the skills variable to test the relationship between skills and technology:

(a) same as existing specification *YRKC*YRKS* (with different definition of skills);

(b) as a separate explanatory variable, i.e. $b_x YRKC + b_Y YRKS + b_z YRKN$.

The existing extra-EU export equation is also used to test the relationship between skills and technology. There are two good reasons for this:

(a) technology cannot be measured. In E3ME the technology terms YRKC are constructed. Therefore the impacts of skills on technology cannot be directly assessed;

(b) exports are highly responsive to technological change. This relationship is well documented in international trade theory (e.g. Barker et al., 2006).

The main drawback of using the export equation to test the link between skills and technology was that it does not provide insight to non-traded sectors such as utilities. A newly developed R&D equation can also be used to test the effect of skills on technology. This is because R&D is highly correlated with new technologies and therefore can be used as a proxy for technology. The R&D equation should be able to give us some insight into the non-traded sectors which the export equation fails to capture.

## 3.4. Involving skills upgrading into the model

Another way to incorporate improvements of technology and skills into the forecasting exercise was tested. It differs from the above in two main aspects: first, different data are used to assess demand shifts between different types of workers and different types of capital; secondly, the main aim is to identify structural breaks at a low level of aggregation, i.e. at the country-sector level. It is based on the work of Acemoglu (2002), who regards technical change as the trigger, whereas institutions, trade liberalisation and organisational change are seen as complementary factors determining the magnitude of the impact of technical change.

Since technology differs across industries and countries, we measure the speed of skills upgrading separately by country and industry. We are interested in the speed at which demand for skills changes over time. Structural breaks occurring in the period 1970-2005 are also analysed. This allows us to recover potential structural changes and identify the direction of the change in the speed of skills-upgrading over time. This analysis is performed separately for countries and sectors as not only the speed but also the timing of the structural break may be dependent on the type of sector and the economic development of the country.

For this purpose, the EU-KLEMS data set is used for nine Member States, Australia, Japan and the US to estimate the speed of skill-upgrading for three skills groups (low, medium and high educational levels), within 11 sectors of industry. The data contain detailed information on labour inputs (low-, medium-, and high-skilled labour), capital inputs (IT and non-IT related capital) and intermediate inputs. These variables are collected at sector level for the period 1970-2005. This information enables the estimation of the parameters of a production function at sector level within each country, allowing for sector- and country-specific production parameters.

Section 3.4.1 describes the theory and empirics underlying the estimation of the skill composition of the workforce. An overview of the data is also given in Section 3.4.1. Results are then presented followed by conclusions.

### 3.4.1. Modelling skill upgrading

The analysis of skill upgrading requires estimating the demand for skills. We do this by formulating a parametric cost function that captures (changes in) technology used within industries and countries. While various parametric forms have been used in literature (see Berndt and Christensen, 1973 among others), we start by applying the translog cost function that has the most flexible function form. This functional form can be seen as a second order approximation of the true cost function. An advantageous feature of such a specification is that it nests the constant elasticity of substitution cost function and the Cobb-Douglas cost function as special cases ([17]).

Let gross output *Y* in one industry of a country at time *t* be produced with *K*=6 different inputs: three types of labour (low-, medium-, high-skilled), two types of capital (information and communication technology (ICT) related capital and non-ICT related capital), and intermediate inputs. Considering heterogeneous labour and capital inputs allows us to test, for instance, how far the use of ICT technology substitutes for low-skilled labour.

Based on the translog cost function, share equations of each input can be derived. Using the traditional stability restrictions on the parameters of the cost function, this system of *K* equations can be reduced to *K-1* equations without loss of generality (with the *K-th* input factor being excluded). For each of the *K-1* inputs, the share equation can be written as (we drop the indexes for industry, country and time for notational simplicity):

$$S_i = \alpha_i + \sum_{j=1}^{K-1} \gamma_{ij} \ln\left(\frac{P_j}{P_K}\right) + \gamma_{iY} lnY + \gamma_{iT} TR \qquad (4)$$

([17]) We have also experimented with other flexible cost functions and, in particular, the generalised Leontief shape. While empirical estimation of the magnitude and sign of the Allen elasticity of substitution (AES) gave similar results for the translog cost function and the generalised Leontief cost function, the magnitude and sign of the estimates of the trend parameters (upon which the speed of skill upgrading depends) using the generalised Leontief function were unrealistic.

Where $P_i$ is the price for a unit of input factor *i*, *Y* is total output, and the variable *TR* is a linear time trend. The dependent variable $S_i$ is defined as the costs of particular input (also ICT) relative to total costs. For any input *j*, it is defined as:

$$S_j = \frac{P_j X_j}{C} \qquad (5)$$

Where $P_i$ is the price of input *j*, $X_j$ its quantity, and *C* the total costs of production, i.e. the sum of costs of all labour, capital and intermediate inputs.

The model of *K-1* equations can be estimated using 'seemingly unrelated regression' (SUR) techniques (Zellner, 1962). This technique allows us to model the contemporaneous correlation between idiosyncratic shocks in the demand for each input ([18]).

The model is estimated for each country-industry time series of output and inputs. This leads to 132 separate regressions (11 industries x 12 countries). Based on the estimated regression coefficients, we can calculate the own-price elasticities $\left(\sigma_{ii}^{AES}\right)$ and the elasticities of substitution between inputs *i* and *j* $\left(\sigma_{ij}^{AES}\right)$. We use the Allen elasticity of substitution (AES) which is defined as:

$$\left(\sigma_{ii}^{AES}\right) = \frac{\gamma_{ii}+S_i^2-S_i}{S_i^2} \text{ and } \left(\sigma_{ij}^{AES}\right) = \frac{\gamma_{ij}+S_i-S_j}{S_i-S_j} \qquad (6)$$

The time trend variable *TR* captures trends in the demand for inputs over time and therefore accounts for shifts in demand which are unrelated to changes in the level of other inputs. This type of skill upgrading would therefore occur if the coefficient is larger for skilled labour than for unskilled labour. The speed of skill upgrading will depend on the relative magnitude of these coefficients across skill groups. A change in the speed of skill upgrading would occur if the relative magnitude of the time trend coefficients across skills groups changes over time. We test for changes in the speed of skill upgrading by testing for structural breaks in the time trend coefficients.

$$speed_{skillupgrading_j} = \gamma_j - \gamma_{LS} \text{ for } j = HS, MS \qquad (7)$$

### 3.4.2. Data

Data from the EU-KLEMS project are used which contain information on growth accounts from 28 European and non-European countries (Timmer et al., 2007).

([18]) However, if current shocks in the demand for each input are correlated with past shocks in the demand for other inputs (i.e. endogeneity of prices), seemingly unrelated regression (SUR) estimates will be inconsistent. To account for this problem, one could attempt to apply instrumental variable techniques using internal instrumental variables as suggested by Krusell et al. (2000).

The data consider 21 industries over a maximum period from 1970 to 2005. They contain input measures on capital (differentiated by information technology-related capital (ICT) and non-information technology-related capital (NICT)), labour (differentiated by low-, medium-, and high-skilled labour groups), and intermediate inputs (differentiated by energy, material, and service inputs).

Not all countries were suitable for inclusion in the estimation sample. Some countries were not included because the span of observations was too short or because important information to generate the required variables (input and prices) was missing. Only nine Member States (Belgium, Denmark, Spain, France, Italy, the Netherlands, Austria, Finland and the UK) and three non-member countries (Australia, Japan and the US) provided satisfactory data which were adequate for running the regressions (more in Dupuy and Sauermann, 2011).

### 3.4.3. Implication for the forecasting model

Over the last decades the skill composition of labour inputs has changed substantially. The general pattern is that the relative importance of low-skilled labour has decreased while the share of high-skilled labour has increased. Figure 4, based on the EU-KLEMS data used for this exercise, shows the share of low-, medium-, and high-skilled labour inputs in total labour inputs. The input of labour is measured as number of hours worked and can be split according to the contribution of each of the three labour inputs. Since all three shares sum up to unity, the lines shown in Figure 4 clearly indicate that the share of low-skilled labour decreased in all sectors over the period 1970 to 2005. While medium-skilled labour remained at a constant level, the share of high-skilled labour substantially increased over this period.

The main argument put forward in economic literature is that skill-biased technological change (SBTC) led to a change in the skill structure. The idea is that new technologies require higher levels of skills for two reasons: first to develop new technologies and secondly to apply them.

Due to SBTC, information and communication technologies (ICT) have been on the rise in recent decades compared with non-ICT technologies. The trend towards increased use of ICT is shown in Figure 5, which depicts the share of ICT capital in total capital inputs over the period 1970 to 2005. In all sectors, the share of ICT-capital in total capital investments rose from almost 0 to very high levels. As we will see below in the estimation of the parameters of substitution (the $\gamma_{ij}$ parameters), this dramatic change in the share of ICT makes it difficult to accurately estimate the structure of substitution.

Figure 4. **Shares of low-, medium-, and high-skilled labour in total labour by industry**

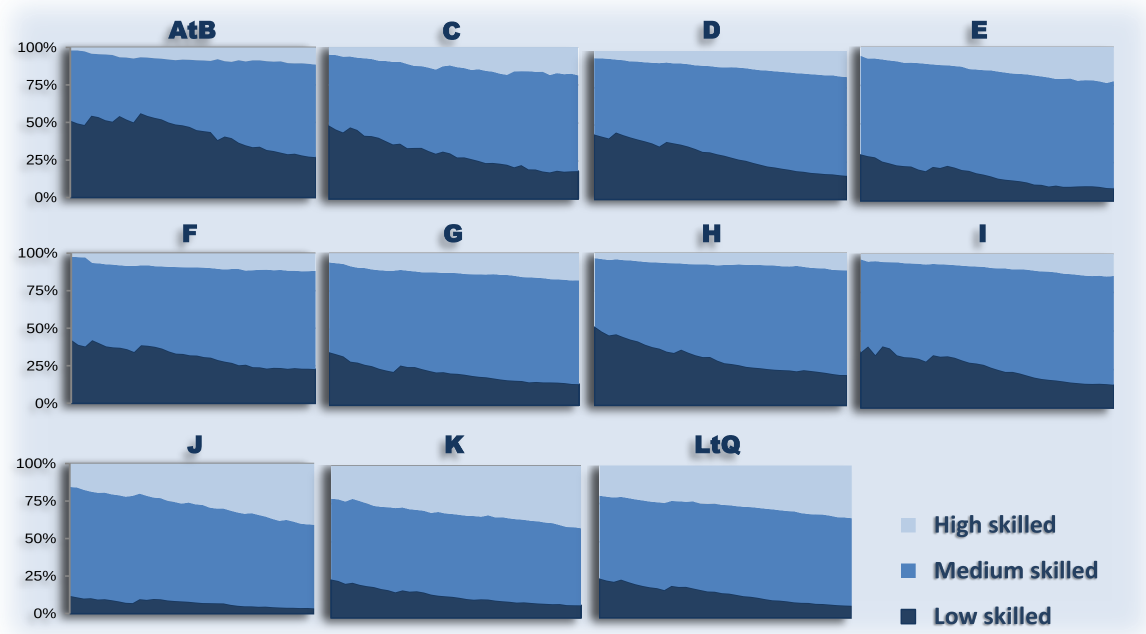
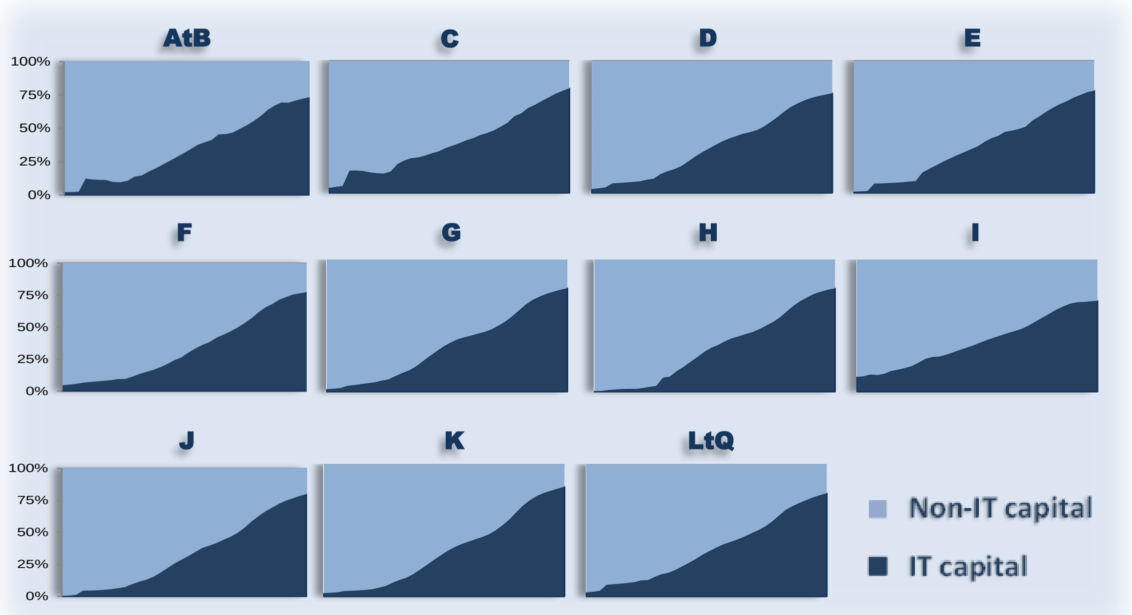



Figure 5. **Shares of IT and non-IT capital inputs in total capital inputs**




NB: Sectors are agriculture, hunting, forestry and fishing (AtB), mining and quarrying (C), total manufacturing (D), electricity, gas and water supply (E), construction (F), wholesale and retail trade (G), hotels and restaurants (H), transport and storage and communication (I), financial intermediation (J), real estate, renting and business activities (K), community social and personal services (LtQ)

*Source*: EU-KLEMS; calculations by ROA.

From a theoretical point of view, the reason for this can be easily seen from the formula of own-price elasticity:

$$\left(\sigma_{ii}^{AES}\right) = \frac{\gamma_{ii}+S_i^2-S_i}{S_i^2} \quad (8)$$

If the cost share of input *i* goes towards 0, the own-price Allen elasticity of substitution (AES) tends to $+\infty$ for positive values of $\gamma_{ii}$ and $-\infty$ for negative values of $\gamma_{ii}$. However, when the share of the input is equal to 0.5, the own-price elasticity will be negative as long as the parameter $\gamma_{ii}$ is not too high, i.e. $\leq 0.25$. This means that, using data where the share of an input *i* changes dramatically from values close to 0 to values close to 0.5, the point estimate of the $\gamma_{ii}$ parameter should not be too high (either negative or positive but close to 0). While at the end of the series, when the share is close to 0.5, the elasticity is negative even for positive values of $\gamma_{ii}$. At the beginning of the series, when the share is close to 0, positive values of $\gamma_{ii}$ would produce positive own-price elasticities.

The estimates of the elasticities tend to have large magnitudes compared to what economic theory would predict. From an empirical point of view, there does not seem to be enough variation around the trend for most price series to avoid multicolinearity problems in the estimation procedure. However, even though the magnitudes of the elasticities are sometimes unrealistic (multicolinearity), their signs appear to coincide with economic intuition. For instance, most of the own-price elasticities are negative (Table 6). This seems to indicate that although the parameters of substitution suffer from multicolinearity and should be interpreted as reduced form parameters, the time trend coefficients which are the main interest to enter E3ME model, are estimated consistently. This technique is possible use in the whole forecasting exercise.

Table 6. **Share of negative own-price Allen elasticity of substitution (AES)**

| Input | Share negative AES (%) |
|---|---|
| ls | 58.0 |
| ms | 73.3 |
| hs | 55.7 |
| it | 85.5 |
| nit | 51.9 |
| ii | 96.9 |

NB: ls – low-skilled; ms – medium-skilled; hs – high-skilled; it – IT capital; nit – non-IT capital; ii – intermediate input.

*Source*: EU-KLEMS; calculations by ROA.

*Baseline results*

The estimated coefficients from the translog cost function allow us to draw conclusions on trends in the demand for low-, medium- and high-skilled workers as well as for ICT and non-ICT capital (for detailed numerical results see Dupuy and Sauermann, 2011).

Loosely speaking, these demand shifts are believed to capture the impact of technical change. As expected, the demand for low-skilled workers is decreasing in most countries and most industries whereas the demand for high-skilled labour is increasing during the same period.

Interestingly, in the agricultural, hunting, forestry and fishing industry (AtB), mining and quarrying industry (C) and manufacturing industry (D) the demand for low-skilled workers decreases consistently in all countries, although with various magnitudes. However, for some industries such as wholesale and retail trade, the demand for low-skilled workers increases in some countries, i.e. Australia and the US.

Regarding the demand for high-skilled workers, while most estimates are positive as expected, the strongest shifts in the demand for high-skilled workers occurs in the following industries: hotels and restaurants, community social and personal services, and real estate, renting and business activities.

Based on these estimates, we can calculate the speed of skill upgrading by taking the difference in the percentage changes of the demand for medium- or high-skilled labour to that of low-skilled labour. The relative demand for medium- and high-skilled workers increased over time although there are large variations in the magnitude across countries and industries.

In manufacturing industry, for instance, the speed of skill upgrading, as measured by the difference between the demand for high-skilled workers and low-skilled workers, tends to be uniform in all countries and lies between 0.01% in Italy to 0.66% in Belgium. In contrast, the magnitude of skills upgrading in the sector of wholesale and retail trade varies among countries, for example standing at 1.41% in Australia and 1.43% in Belgium.

CHAPTER 4.
# Modelling skills supply

The methods currently used to produce the skill supply forecast were developed in the first pilot project of Cedefop (2009) and improved in the following stages of the research (Livanos and Wilson, 2010a; Bosworth and Wilson, 2011).

Although there has been a great deal of academic work on educational choice and related issues, it has only rarely focused on the development of models suitable for developing projections of the supply of skills. Some work in individual Member States has made progress in this direction, but this has usually been based on access to specific data not generally available at pan-European level.

Best practice involves a full analysis of both stocks and flows. Stocks in one period are related to stocks in an earlier period, plus inflows and less any outflows. The flows can be linked to demographic developments and to a range of behavioural drivers, including economic and social factors. Livanos and Wilson (2008) briefly review this research before presenting a much simpler approach, focusing solely on changes in the average qualification patterns within stocks, which reflects the limitations of the data available at pan-European level.

The current methodology is based on analysis of changing patterns over time in the stocks of people in the population and in the labour force, defined by highest qualification held, and by country separately. These data are taken from the EU-LFS. However, it is clear from the analysis that there are still some problems of comparability among countries due to how qualifications are coded within the LFS, as well as difficulties in some countries in comparisons over time as a result of unrecorded discontinuities due to changes in classification or definitions used. Nevertheless, the results suggest that even with such imperfect historical data on the highest qualifications attained, it is possible to begin to explore the implications for the future supply of qualifications in the population and the labour force. Many of the trends appear to be very robust and common in all countries.

Ideally, in the longer term, this approach would be extended to include a much more sophisticated analysis of the supply side, including a greater focus on the factors influencing flows, and greater detail on the types of qualifications held.

Another important issue is the relationship between skill supply and demand. The present approach focuses on changing supply patterns, without any reference to demand-side developments. In practice, the observed changes are likely to have been the result of a combination of both demand and supply influences.

As a measure of skill supply, the highest formal qualification obtained by those who entered the labour force is used, disaggregated by five-year age groups and gender. In general, skills supply is determined by demographic developments, labour market participation and decisions on obtaining an educational credential. As mentioned in Chapter 3, the calculations determining the volume of labour supply projections are developed within the framework of the E3ME model. This chapter is related more to the forecasting of the qualification patterns.

Section 4.1 presents the key features of the stock model for supply by qualification currently used. Section 4.2 describes the theoretical background of stock-flow model.

## 4.1. Specification of stock model for supply by qualification

There are various theoretical considerations surrounding the modelling of skills supply, including factors that might explain changes in the qualifications structure of the labour force and of the population as a whole. However, the estimation of complex behavioural models is rather complicated in practice, mainly due to data restrictions.

Regarding modelling educational choices at individual level, the available data rarely contain the socioeconomic background information that is necessary for such modelling. Moreover, there is no information on abilities and other factors that are known to the individuals at the time they make their decision regarding educational attainment. These factors influence the level of education but unfortunately are not available in the LFS or other data sets, and thus any analysis that lacks this information might produce biased results. Similarly, there is no information on wages so the impact of economic factors cannot be estimated easily.

Regarding modelling at aggregate level, there is scope for including some behavioural elements in future stages of this project to help us to understand factors affecting educational attainment. In particular, the aggregate data from all countries can be pooled together creating a unique data set. To this data set various explanatory variables can be considered. For instance, one variable that in theory may measure a potential motive for educational attainment is the unemployment rate broken down by age, gender and qualification level. Other possible variables that can be included in the modes are expenditure on education as a percentage of GDP and financial aid to students defined as a percentage of total public expenditure on education. These data are available at country level, and can help capture common patterns and differences among countries. Including these variables in the model could add some behavioural insight.

Livanos and Wilson (2008) reviewed the theoretical and empirical approaches to modelling and projecting the supply of people with different levels of qualifications. Where detailed data are available, researchers have exploited them. More often than not, however, the paucity of information available has resulted in very simple approaches based on time series methods, which themselves have often relied on a single variable (time), rather than on multivariate, behavioural approaches.

The ideal approach would probably be to use a linked stock-flow model, with future stocks being related to past ones by an accounting relationship, including separate analysis of all relevant inflows and outflows. The latter would include flows of people into education and of those obtaining formal qualifications, as well as outflows for various reasons (including mortality, migration and flows in and out of economic activity).

Given suitable data, such flows can be explained using behavioural models, such as the human capital approach. The latter argues that education can be regarded as an investment and that the decisions of people to participate in education and to acquire qualifications are influenced by the returns to such investment. Based on such theories the supply of skills (as measured by qualifications) can be related to career choices influenced by a range of economic and other determinants, including:

(a) pay (returns to investment in qualification),
(b) employment opportunities when searching for a job,
(c) social security and tax system,
(d) personal and household characteristics,
(e) other economic and social factors.

Previous research on these issues at pan-European level is quite limited, especially at detailed level. In practice, the data to build such models are not generally available at pan-European level. The length of time series of the data available is also very short. In practice, such data only allow for fairly very simple extrapolative procedures (using linear or non-linear methods). Effectively, time is the only independent variable, acting as a proxy for other factors. This creates a large gap between the ideal theoretical model and the typical specifications used in most national level skills supply projections.

To assess changing patterns of skill supply within each country, a simple model for developing projections of the supply of people by qualification has been developed in the Cedefop project. This model focuses upon patterns in the qualifications currently held as reported in the LFS. These patterns are then applied to independent projections of the population aged 15+ and to the labour force, each differentiated by age and gender as taken from the E3ME model.

The propensity to hold a given level of qualification is measured either using aggregate data or at individual level (the latter using a pooled cross-sectional data set drawn from the LFS). The trends are estimated using a multinomial logit model or something similar. Variables are introduced into the model to specifically capture ‘spatial effects’ (including country specific intercept and time trend for each qualification level, age and gender category). This modelling strategy makes maximum use of the limited sample size in the LFS. The analysis focuses on individual qualification achievement levels (dependent variable) against a set of explanatory variables. Demographics are controlled for by including age-band and gender categories between which qualifications may be expected to vary.

The regression analysis of individual data uses a multinomial logit specification. The propensity to obtain a particular level of qualification at the individual level is given by:

$$Q_{ijk} = \alpha_{jk} + \beta_{jk}t + \gamma A_{ijk} + \varphi G_{ijk} \qquad (9)$$

where, *i* denotes the *i-th* individual, *j* refers to the *j-th* country and *k* is the *k-th* age group. *Q* takes values of 0 or 1 depending on whether the individual holds a particular qualification or not. There are three qualification categories: high, medium and low (one category is left out as the base group). *A* denotes a set of age-band dummies, taking values of 1 if the individual falls into a particular age category. *G* is a gender dummy, taking values of 1 for females and 0 for males. Finally, a set of time trend variables *t* are included, capturing changes in attainment levels over time which differ by country and gender and for each age band. The model is estimated using the multilogit procedure in the STATA software package, which constrains the proportion of individuals holding each qualification level to sum to unity.

The most robust results are obtained for some of the simplest models. These generally involve some form of simple time trend rather than more sophisticated models with behavioural content. Given the problems with the current data this is probably not surprising. As the data are extended and improved it may be possible in future work to add in more economic and other behavioural content to this part of the modelling.

The projections of working age population (and the labour force) by highest level of qualification held can be obtained by applying the projected probabilities of the logistic specification to the working age population (and the labour force) projections from Cambridge Econometrics and the E3ME model ([19]). For a discussion of the underlying methodology used to develop the projections of the working age population and the economically active labour force see Pollitt and Chewpreecha (2009).

The projected shares of educational attainment by gender and age groups are applied to the projections of the working age population for the same age and gender categories. This generates forecasts of the numbers of people in the working age population by highest qualification held, broken down by age and gender. The projections of the labour force are obtained by applying the estimated activity rates to the projections of the working age population.

The role of individual country experts is of considerable importance for the project, since the results obtained need to be validated using local knowledge. In particular, the country results obtained from this stage of the project are typically examined by individual country experts to have their feedback incorporated into the final or updated set of results.

## 4.2. Building a stock-flow model

The possibility to develop more sophisticated analyses using information on the flows of people through the education system has been investigated. The overall supply of people holding formal qualifications at higher level (ISCED 5-6) is relatively straightforward to conceptualise and model. However, there are considerable conceptual and practical difficulties in extending this to lower levels of education (ISCED 1-2). The main difference of these two groups is that lower qualified may increase their educational level during the working life and so initial stock may be changed in future. Conceptualising the idea of supply to cover specific dimensions such as occupation, sector and geographical area is even more difficult. This is primarily because the educational systems in most countries are not completely hierarchical (Cörvers et al., 2008).

The key problem to be addressed in extending stock-flow models to cover the full range of qualifications is the much more limited information available at lower level qualifications (ISCED 1-2). Ideally, stock-flow modelling requires a

([19]) The population projections that Cambridge Econometrics uses to produce forecast of the labour force is drawn from Eurostat.

comprehensive set of demographic accounts showing how individuals progress through the educational system and the labour market over time. In practice such accounts do not exist, although there is a considerable amount of information on certain flows.

If educational transitions stopped at an early age (e.g. 22), it might be possible to assume that the qualification mix would stabilise for groups older than this. Under these circumstances, a stock-flow model would stand a chance of being a reasonably reliable method of saying something about the evolution of qualification levels into the future.

#### 4.2.1. Essence of a stock-flow model

The first element of the stock-flow model is the knowledge that a cohort of individuals identified at time *t* can be largely traced at time *t+1*. In other words, a group of 16 year-olds in 2007 become 17 in 2008 and so on. The example is provided in Table 7. The starting 'stock' comprises individuals of working age (16-64) in 2007. Inflows are needed at the youngest age level ([20]) (16) to fill the shaded area in the top right hand corner, as the stock ages. This is normally possible in demographic models because there are usually data on 15 year-olds in 2007, who become 16 in 2008 and 17 in 2009.

Table 7. **'Rolling the LFS data forward' in a stock-flow model, the UK**

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | ... | 2020 |
|---|---|---|---|---|---|---|---|---|
| 16 | 1 094 678 | | | | | | | |
| 17 | 985 209 | 1 094 678 | | | | | | |
| 18 | 900 975 | 985 209 | 1 094 678 | | | | | |
| 19 | 843 577 | 900 975 | 985 209 | 1 094 678 | | | | |
| 20 | 884 830 | 843 577 | 900 975 | 985 209 | 1 094 678 | | | |
| 21 | 912 422 | 884 830 | 843 577 | 900 975 | 985 209 | 1 094 678 | | |
| 22 | 926 265 | 912 422 | 884 830 | 843 577 | 900 975 | 985 209 | | |
| 23 | 883 788 | 926 265 | 912 422 | 884 830 | 843 577 | 900 975 | | |
| 24 | 882 461 | 883 788 | 926 265 | 912 422 | 884 830 | 843 577 | | |

NB: Data for the UK population as a whole, 2007.
*Source*: EU-LFS.

Assumptions also have to be made about deaths and migration from intermediate ages of the stock (e.g. deaths among 50 year-olds in 2007 or net

([20]) Sometimes referred to as the 'intensive margin' in vintage modelling.

migration among 30 year-olds in 2007). Normally such factors are accounted for by simply scaling the LFS 'projections' up to the population forecasts provided by the appropriate bodies for each country.

In general, such scaling of the data to projected population numbers only takes place after the main elements of the modelling – in the present case, educational mix – have been carried out. Thus, attention turns to modelling education and educational transitions.

#### 4.2.2. Qualifications in a stock-flow setting

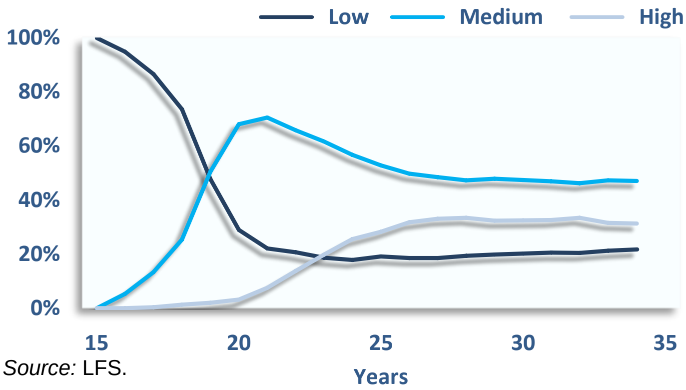
Given that the qualification levels of the population are known in 2007, then the stock-flow model can be formed by qualification level in the same way as for the population as a whole. The 'stock' part of the stock-flow model is primarily useful for those ages above which the qualification levels of each member of the population effectively become fixed. At this point, the mix of qualifications is constant.

Thus, the 'stock' part of the stock-flow model focuses on qualification proportions. It is immediately clear that the proportions shifted forward one year from 2007 to 2008 are unlikely to reflect the true proportions of the qualification mix, at least for the younger ages. The problem is that individuals above the age of 16 undergo educational courses that eventually lead them out of a low qualification level into the medium category and, for slightly older individuals, out of medium into high. If the LFS data for younger individuals is to be used at all, something needs to be done to adjust the data for educational transitions, at least among the younger individuals, as the 2007 cross-section is 'rolled on' for 2008, 2009, etc.

#### 4.2.3. Educational transitions and trends: evidence from cross-sectional data

Educational transitions are reflected in any annual cross-sectional data for individuals, for example, the stock of individuals of working age in 2007, as shown for the EU-27 in Figure 6. The downward shift in low qualifications and the corresponding upward jump in medium qualifications between the ages 16 and 20 are mainly the reflection of the educational transition that occurs every year in the EU as individuals obtain their secondary education qualifications. Likewise, the downward shift in the medium qualification proportion and the upward jump in the high proportion around ages 21 to 28 reflect the educational transition that, again, broadly speaking, occurs every year in the EU as individuals obtain their tertiary qualifications. It is clear that similar patterns could be recognised in each Member State. They may of course vary with intensity and over time.

Figure 6. **Qualification proportions of a cross-section of individuals, EU-27, 2009**




*Source:* LFS.

The curves in Figures 6 reflect two effects: education transitions, which occur from year to year and educational trends, which affect the size (and timing) of the transitions in the past. As educational transitions are particularly important in the early years, they tend to dominate among younger age groups in the cross-sectional data and, as educational transitions become less important among older individuals, the long-term educational trends dominate the curves among the higher age groups. An age factor is also possible to observe as there is rising proportion of older individuals with low level qualifications and the falling proportion of those with high qualifications. This is based on fact that a smaller proportion of earlier population cohorts went on to higher education.

### 4.2.4. Constructing pseudo-cohorts from the LFS data

*General principles of pseudo-cohort modelling*

The entry level of a given cohort and the changes in qualifications for that cohort as individuals age is called a ‘within cohort’ effect, while the differences at entry (or over time from one cohort to another) is called a ‘between cohort’ effect. Both effects need to be modelled, and this is where the pseudo cohort data by level of qualification attainment become potentially important.

The pseudo cohort data are a potential mechanism for modelling the transitions between educational attainment levels and filling in the missing qualification information about the inflows of young people of working age that characterise forecasts based upon stock-flow modelling.

Using a pseudo-cohort constructed from the EU-LFS, it is possible to observe individuals of a given age in 1995 over the succeeding 13 years (i.e. 1995 to 2007). This is a pseudo-cohort and not a real one. It is a comparison of largely different individuals aged, say 16 in 1995, 17 in 1996 and so on.

Nevertheless, if, for example, the group of individuals age 16 in 1995 is representative of the relevant population group and the same is true of those aged 17 in 1996, it is possible to treat them as part of a cohort. By implication, individuals aged 17 in 1996 will have been through the same education experience up to 1995 as those individuals aged 16 in 1995, and the same holds for those aged 18 in 1997.

*Operationalising the pseudo-cohorts*

The way in which the data are organised to obtain pseudo-cohorts is shown in Table 8. The left hand column gives the year of entry of that cohort into the population of working age (assumed to be 16); the first row gives the subsequent age of that cohort; and the number within each cell is the year of the EU-LFS from which the data for that cohort at that age are obtained. Since at the time of this analysis the most recent data were of those from 2007, only the first year for the entry cohort of 2007 was available. As 1995 is the first year of data available, it is not possible to observe 16 year-olds in 1994 (only 17 year-olds in 1995). Thus, all of the shaded cells in Table 8 cannot be observed.

Table 8. **Year of the EU-LFS used in constructing pseudo cohorts**

| Year of entry (age 16) | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2007 | 2007 | | | | | | | | | | | | | |
| 2006 | 2006 | 2007 | | | | | | | | | | | | |
| 2005 | 2005 | 2006 | 2007 | | | | | | | | | | | |
| 2004 | 2004 | 2005 | 2006 | 2007 | | | | | | | | | | |
| 2003 | 2003 | 2004 | 2005 | 2006 | 2007 | | | | | | | | | |
| 2002 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | | | | | | | |
| 2001 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | | | | | | |
| 2000 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | | | | | |
| 1999 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | | | | |
| 1998 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | | | |
| 1997 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | | |
| 1996 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | | |
| 1995 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | |
| 1994 | | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |

*Source*: Kriechel (2011).

If the data related to a true (as opposed to a pseudo) cohort, there would be an expectation that the low qualification category would monotonically fall and the high qualification category would exhibit a monotonic rise with the age of the cohort, at least up to some point at which each of them become constant.

*Shifts in the historical pseudo-cohorts*

The basic hypothesis in the data analysis is that if the patterns of the shares of particular qualification levels remain unchanged across different pseudo cohorts, this might imply that the same pattern could be imposed to the new inflows of young individuals through to 2020. The analyses of data for the UK confirmed that in practice this does not appear to be the case (Bosworth and Wilson, 2011). There appears, for example, to be a marked change among the low qualification group in the UK pseudo cohorts for 1995-97 in particular, and in subsequent cohorts. These findings raise the crucial question of how to model the transitions through to 2020, to enable the stock-flow approach to function adequately.

*The need to model the effects of migration on qualification mix*

An important issue in the modelling process is the skills lost through emigration and gained through immigration, although there are a number of problems associated with this. Then the question is how the stock-flow model estimates of qualifications should be adjusted for migration activity. While in some countries migration may be a minor issue, in others it may be much more important. In addition, while in some countries the qualifications of the inflow may be very similar to that of the outflow, in others the two flows may involve very different qualification mixes. Cedefop attempts to work on the estimation of qualification mixes of migrants. For more information on this consult Gausas et al. (2011).

It seems likely that much more could be done with the EU-LFS data in modelling the effects of migration on the qualification mix of each country, at least at the three main levels of qualification. This might be carried out by modelling the historical migration transitions ([21]) by qualification level and pinning the results to the Eurostat total migration information, by age and gender. The Eurostat data also give the country of origin of the immigrant population.

([21]) Flows might be modelled in various ways, e.g. using 'years of residence in this Member State (YEARESID)' or 'country of birth (COUNTRYB)' (Eurostat, 2008).

*Linking the pseudo-cohort modelling with educational flows data*

In addition to the analysis of the pseudo-cohort data, there is a strong case for linking these educational transitions results to direct evidence based on educational flow data. In particular, the analysis of aggregate published Unesco/OECD/Eurostat (UOE) flows data (on enrolment and graduation by ISCED level) can be used to produce a complementary analysis of participation and qualification rates by broad age groups.

These data can be used both in the context of a historical analysis of educational transitions and in a forecasting context. They might be the subject of time-series modelling that might throw light on various issues, such as the minimum proportion of low qualifications that can be achieved by each Member State.

CHAPTER 5.
# Modelling employment growth (expansion demand)

Research on modelling employment by occupation and qualification and the development of relevant Cedefop module is summarised in Livanos and Wilson (2007b, 2010b). This method was used in the pilot project *Future skill needs in Europe* (Cedefop, 2008) and in *Skill supply and demand in Europe* (Cedefop, 2010). Latest developments are summarised in Bosworth and Wilson (2011).

This chapter describes the procedures to estimate the occupational shares. It illustrates various techniques and data sets explored to improve module 2 (EDMOD) and module 3 (QMOD) of the Cedefop modelling framework (Figure 1). Labour demand within sectors is the scope of the E3ME model (Chapter 3).

The chapter reviews recent literature on occupational modelling as well as the specifications and data used in previous research to examine how the data available can facilitate the development of a ‘best practice’ approach at pan-European level. In addition to presenting various simple extrapolation approaches, the chapter explores a new method proposed for estimating occupational shares using a multinomial logistic regression, rather than the simple extrapolation of past trends. The same method is also used to analyse qualification shares within occupations.

The first section focuses on relevant data specification of an ideal model. The second part is dedicated to specifications of the current model. Final part of this chapter describes the testing of other solutions and approaches.

## 5.1. Data available and the ideal specification

Most occupational forecasts are based on either fixed-share coefficients or simple trend-extrapolated coefficients. For example, in Cyprus (Oxinos et al., 2005) forecasts by occupation are obtained by applying past occupational shares to future data of sectoral employment. In Ireland (Hughes and Fox, 2005) the occupational structure of employment is projected by taking account both of past trends and expectations relating to the evolution of skills and occupations. Occupational share coefficients for each sector are calculated for the data period, and linear and semi-log trend line regressions of geometric growth rates, where judged appropriate, are used to project the shares for the sectors for the target forecast year.

Similar methods to the above are used in many other EU Member States, and also in countries outside the EU. For example, in Australia the occupational share effects are treated as a type of technical change and are forecast by extrapolating historical trends in the occupational mix in each industry (Meagher et al., 2000; Meagher, 1997). In the United States, the occupational projections of the Bureau of Labor Statistics (BLS) are published in the *Monthly labor review* (e.g. Hecher, 2005). Their overall approach has not changed since the late 1990s (BLS, 1997). Based on this review of literature, the ideal specification for modelling occupational employment might be a behavioural model as shown in equation 9:

$$S_{ijt} = f\big(Year, Technology_{ijt}, Trade_{ijt}, Wage_{ijt}, Output_{ijt}, Output_{ijt}, Unemployment, X\big) \quad (10)$$

where *S* denotes the share of employment by occupation *j* within industry *i*, and *X* denotes a vector of other employment characteristics.

This represents the long-term relationship between the demand for skills and various key drivers. In practice it can be anticipated that there will be various lags and adjustments which require the exploration of the dynamics of this relationship.

Finding suitable data to test such a model is difficult, even at individual country level. The data that is considered for the present occupational modelling exercise come mainly from two main sources.

First, the project utilises aggregate pan-European data from the EU-LFS, as published by Eurostat. These data cover all 27 Member States for the period 1993-2005, and contain information on employment by country, gender, industry (41 industries), occupation (27 occupations) and qualification (3 level qualifications). These data measure the changing pattern of skill demand ([22]) and are in theory harmonised to a common classification.

To explore the possibility of investigating the importance of various economic indicators on the occupational structure, economic indicators by sector from the E3ME model were included in the data set. These were included to try to find links between changing economic drivers and the pattern of employment by skills. The E3ME data set includes many possible indicators by industry such as gross value added (GVA), hours worked, average earnings, gross output, imports and exports, as well as technological indicators such as expenditure to research

([22]) However, the observed employment levels are the consequence of both demand and supply factors. The latter may be especially important with regard to the qualification dimension.

and development (R&D) and investment in information and communication technologies (ICT). When combined the Eurostat and E3ME data allow for, in principle, the estimation of a detailed specification such the one outlined in equation 9, which includes the impact of the business cycle and of exposure to international trade. However, various technical and data problems (e.g. short time series) restrict the ability to allow for simultaneous analysis of all industries.

The second main data set used is the LFS microdata set on individuals. These data, covering the period 1983-2005, provide useful labour market information at individual level. They provide a great deal of information on demographic characteristics (age, gender, nationality, etc.); an individual's labour market activity over the reference week (employed/unemployed, etc.); characteristics of an individual's first job (occupation, sector, employment status, hours of work, etc.); flexible working patterns (evening work, overnight, etc.); second job; information on an individual's previous employment (if unemployed); methods of looking for a job; and education (for a full description of variables included in the LFS see European Communities, 2003).

## 5.2. Current specifications of the model

Given the problems with the data, and while the ideal specification might look like equation 9, several simpler specifications have also been considered. These include a range of models of the general form shown in equation 10.

$$S_{ijt} = F(time) \qquad (11)$$

Two main methods of analysis were adopted: analysis of the employment shares in the published LFS data and adopting specifications as in equation 10. These include several variations including linear, semi-log and logistic forms. Such equations can be regarded as simplified versions of equation 9, where time is used as a proxy for technological change.

Several variants are possible. These range from simple extrapolation between fixed points to various methods based on 'line fitting'. The latter includes fitting:

(a) a linear trend,

$$[S = a + b * Time] \qquad (12)$$

(b) a log linear trend

$$[ln(S) = a + b * Time] \qquad (13)$$

(c) or a logistic equation

$$\left[Ln\left(S/(1 - S)\right) = a + b * Time\right]$$

(14)

Another method of analysing changes in the employment structure involves a probability model, estimated on individual or pseudo-individual data. The propensity to be employed in a particular occupation is modelled using a multinomial logistic regression framework, based on pooled cross-sectional data from the EU-LFS data ([23]).

A multinomial logistic regression model can be used to estimate the probability of an individual working in occupation (OCC) *j* at time *t*. The general model is specified as follows:

$$\Pr(OCC = j|T = t) = \frac{\exp\left(\mathbf{\Omega}^{(j)}X\right)}{1 + \sum_{i=1}^{N} \exp\left(\mathbf{\Omega}^{(i)}X\right)} \tag{15}$$

and for an arbitrarily chosen base category:

$$\Pr(OCC = N|T = t) = \frac{1}{1 + \sum_{i=1}^{N} \exp\left(\mathbf{\Omega}^{(i)}X\right)} \tag{16}$$

The equations state that the probability of the representative individual working in occupation *j* at time *t* can be expressed as a function of explanatory variables, normalised by the sum of probabilities for all categories. There are *N* occupations and the sum of probabilities is constrained to add up to one (i.e. shares of employment across occupations sum to 100%).

Similarly, the above model can be used to estimate the probability of an individual working in occupation (OCC) *j* at time *t* holding a qualification *n*.

In the model *X* relates to a vector of regressors which are included in the model as explanatory variables for occupational or qualification structure. Potentially these might include:

(a) wage of occupation *j*,

([23]) For the purposes of the analysis both micro-LFS and a pseudo-individual data set (based on published LFS) were used. However, due to lack of consistent and reliable data in the final results the microdata set was rejected. The pseudo-individual data set based on the published Eurostat numbers was used for the mlogit analysis reported here. This was created by generating large numbers of pseudo individual cases with the same properties (occupational and qualification shares) as the published estimates.

(b) unemployment rate of occupation *j*,
(c) sectoral value added,
(d) export volume share (exports as a percentage of domestic value-added),
(e) import volume share (imports as a percentage of domestic value-added),
(f) sector of employment (41 industry level),
(g) sector of employment *x* time,
(h) country of residence,
(i) country of residence *x* time,
(j) individual characteristics: gender, age,
(k) time period of observation, *t* (LFS quarter and year).

The time trend variable proxies the impact of technology and other factors on changing occupational structure. Time was interacted with country and industry to allow for these effects to be country and sector specific. The regression coefficients are estimated so that the predicted model achieves 'best fit' to the observed data. This is done using the maximum likelihood method. Note that in the model, $\mathbf{\Omega}$ is the matrix of regression coefficients which is used to predict the distribution of employment by occupation at each point in time, *t*. Note that categorical variables (SIC, country, age, gender, etc.) are included in the model as an exhaustive set of dummy variables.

The primary aim of estimating such a model is to provide a behavioural explanation of changing patterns of employment structure. If a robust specification can be estimated, the model can be used to forecast shares of employment forward over time but also indicate sensitivity to a range of economic influences. Using the coefficients of the model, predicted shares of employment by occupation can be obtained.

In practice both the individual data sets (the individual data as well as pseudo data) are very large, comprising large numbers of observations for each country ([24]). This caused serious technical problems when it came to estimation. The initial intention was to estimate general models covering all countries simultaneously. It became apparent that, despite various efforts to get around these problems, this was not technically feasible given the available hardware and software ([25]). Much simpler specifications had to be adopted, focusing on estimation for just one country at a time.

---

([24]) To achieve a reasonable degree of precision, the number of cases created for the pseudo individual data set had to be even larger than the microdata set.

([25]) Estimation was undertaken using the STATA package.

The final analysis was conducted for each industry separately rather than pooling all industries together. Multinomial logit (mlogit) analyses were conducted to estimate the shares of qualifications of each occupation in every industry of each country. The main advantage of the mlogit method is that the sum of probabilities is constrained by the model to add up to one (i.e. shares of employment across occupations/qualifications sum to 100%), whereas the other methods (e.g. linear, simple extrapolation) estimate the share of each occupational/qualification group separately and they are then constrained to add up to unity. Therefore, the mlogit approach ensures a consistent picture across occupations (all probabilities summing to 100%) without having to impose an external constraint.

Despite considerable effort to develop more general models, the most robust results were obtained using simple country data sets, distinguishing occupation and qualification for each sector, based on the published LFS shares and using models that included just ‘time’ as the only independent variable.

For the final set of results an algorithm was developed to select the preferred models to be used for projecting occupational and qualification shares in each sector. This choice depends on the data available and how well the model fits and predicts the shares. The default model is a logistic specification, based on the published LFS shares, with time as the only independent variable, fitted on published LFS data for the period 1993-2006.

Where data are unavailable or inappropriate (due to missing or inconsistently classified information resulting from changes in classification or other discontinuities), the estimation period is truncated accordingly. The algorithm also checks to see if the projected changes are plausible, censoring out shares that lie outside the range zero to unity and also where the projected change is exceptionally rapid (which usually arises as a result of idiosyncrasies in the data). Where there are problems of this kind, the algorithm explores alternative specifications (log liner, liner and fixed shares) until an acceptable outcome is achieved. This applies to both shares of occupations within each sector and shares of qualifications within occupations.

## 5.3. Testing other solutions and approaches

Cedefop is continuously working to improve and fine-tune methods. Recently we have explored the possibilities for refining the treatment of expansion demand using more sophisticated econometric models. The aim was to incorporate more economic drivers and to include other types of analysis that might help to develop additional behavioural explanations of skill changes in Europe. This exercise was mostly done at a theoretical level. The main focus was on using data that are available in the main project database, including the E3ME database, and analysing the standard 41 sectors and 27 occupations ([26]).

Several useful conceptual underpinnings of a 'behavioural' theory of labour demand, relating occupations and education levels, already exist. This theory has its roots in production and cost functions, which have been adapted to estimate employer demand for labour. Subsequent work has been carried out in extending this framework to look at skill-biased technological change (SBTC). In addition, there has been extensive testing of the conceptual elements of literature, which has given support to both theories, namely that there is a capital-high skills complementarity and that there is SBTC in favour of higher level skills and away from lower level skills.

Given the exploratory nature of the present work, not all issues and problems have been resolved. Although many of the results were suggestive and encouraging, the approach has not offered to date a viable immediate alternative to the simpler methods for developing projections of changing skill demand patterns based on extrapolative techniques. Nevertheless, the results obtained do appear to offer some promising areas for future research, if additional resources can be made available to support it.

The databases used for this analysis can be further improved in several ways: first, the inclusion of other E3ME variables (such as approximations for the ICT and non-ICT capital stocks, rather than the corresponding investment variables); secondly, exploration of other Eurostat data that might be relevant to 'behavioural' modelling of labour demands (in particular, whether a measure of R&D spillovers can be constructed along the lines of Scherer, 1982; 1984).

Two particular features of the dependent variable used here create problems:

(a) the number of categories of occupations, coupled with education level (27x3=81);
(b) the fact that the data are grouped (i.e. not individual observations of a 0.1 nature).

A potential rigorous solution to the latter problem has now been identified. This is a module developed by researchers at Boston University which can work within STATA. The maximum likelihood method assumes that the proportions follow a Dirichlet distribution (a multivariate generalisation of the beta distribution

([26]) This exercise was developed for searching to search for possible improvements of the module and currently is currently not a part of the general forecasting framework.

used to estimate fractional logits). The Dirichlet distribution ensures that each education/occupation proportion remains between 0 and 1, and also that all possible education/occupation proportions will add up to 1.

At this stage, however, it is not clear that simultaneous joint estimates of all 81 education/occupation categories could be carried out from a technical perspective. One possible strategy might be to consider the dimensions of the data sequentially (e.g. first at an occupational level and then by level of education). The regressions reported give some indication of the potential of using a two stage procedure, which models occupational demands first and then educational demands.

The initial exploration of the data suggests that interesting patterns emerge in the present results. While these patterns are largely supportive of the skill biased technological change hypothesis – that technological and organisational changes are linked to increased demand for highly skilled individuals – there is also evidence of diversity across occupations and sectors.

While it is probably premature to draw conclusions about the success of further, future work in this area in the context of forecasting, it seems certain that such efforts would add significantly to the understanding of recent and current labour demands and employment outcomes.

The technical details describing the results are available in Bosworth et al. (2011). While this confirms the experience from other countries, which suggests that it is very difficult to establish robust relationships of this nature and that there is only so much that can be done using econometric techniques, it demonstrates that it is important to explore these alternative approaches and to ensure that best practice is being used, given the data available.

# CHAPTER 6.
# Modelling replacement demand

In addition to analysing changes in overall occupational employment levels, it is important to consider replacement needs arising from retirements, net migration, movement into other occupations and in-service mortality. This is referred to as replacement demand. In general terms, replacement demand can be seen as job openings arising because of people leaving the workforce or their occupation. Most work on replacement demand has tended to focus on 'permanent or semi-permanent' withdrawals from the employed workforce. The main reasons for this are retirement, emigration, and especially for women family formation and child-bearing and rearing. Next to withdrawals from the labour force, we also include inter-occupational mobility – movements of workers from one occupation to another. These movements leave a vacancy and hence create a demand for a worker to fill the vacancy.

The concept of replacement demand is used in Cedefop's skill demand forecast since the beginning (Cedefop, 2008). The first section of this chapter provides the general overview of the replacement demand concept. The second section is devoted to cohort component methods ([27]) used to produce the results presented in the latest Cedefop publications (Cedefop, 2010; 2011). The latest attempts to improve the cohort component methods (as in Kriechel, 2011) are reviewed in the last section.

## 6.1. General overview

The main data used to model replacement demand are based on the LFS. These data allows us to analyse the demographic composition of each occupation. It allows us to estimate specific rates of retirement for each occupational class. LFS data can also be used for making estimates of outflow rates. The replacement demand model (RDMOD from Figure 1) has been developed based on similar data sources to the occupational model (ECMOD) (details in Kriechel and Sauermann, 2010). It is driven in part by the occupational and qualification structure of employment, in combination with information on the probability of flowing out from employment due to retirements, occupational mobility and

([27]) Technical details about this method are provided in Kriechel and Sauermann (2010).

migration. The model used builds upon the national model from the Netherlands (Cörvers et al., 2008). Similar models using variants of the methodology are employed in several other countries, both within and outside of the EU.

Projections of occupational employment typically focus on the total numbers of people that are expected to be employed in such jobs in the future. While such estimates can provide a useful indication of areas of change, highlighting the likely net 'gainers' and 'losers', they give a misleading impression of job opportunities and skill requirements. Even where the projections indicate significant employment decline over the medium term, there may nevertheless be quite good career prospects with significant numbers of new job openings. This is because, as long as significant numbers of individuals are still likely to be employed in the future, employers will need to replace those employees who leave for retirement, career moves, mortality, or other reasons. Replacement demand may often dwarf any 'structural demand' or 'expansion demand', resulting from growth in employment in a particular category. It can easily outweigh any negative changes due to projected employment decline.

While the concept of replacement demand is simple enough to grasp, estimating it is a rather different matter. The main problem is that official statistics place much more emphasis on measuring stocks of people in particular states rather than flows, which is essential to estimating replacement demand.

However, use can be made of readily available statistics to provide indicative estimates. Ideally, one requires a full set of demographic accounts that trace people's movement from one socioeconomic position (e.g. employment in a particular occupation) to another (e.g. retirement). In practice, such a complete set of accounts are rare even at national level. However, for several consecutive years now the LFS provides a sufficiently large sample to obtain estimates of the main elements at national level. The key components are: information on the age and gender structure of occupational employment; information on rates of outflows due to retirement (including early retirement); inter-occupational mobility; migration and/or other reasons for leaving the workforce.

The information on outflow rates can also be estimated using stocks of age-cohorts within occupations for several years. Using the year-to-year changes the outflow rates by occupation-age cohort can be estimated. However, these estimates may not allow for discrimination between the reasons for the outflow that leads to replacement demand. Given the data availability, this methodology is used for the purposes of the current forecast.

Availability of following data is considered as useful for the purposes of this methodology:

(a) Age structure.

Data on age structure are required since many of the flows, especially retirements, mortality and occupational mobility, are age specific. Age structures vary significantly by occupation. For some groups, such as corporate managers and administrators, experience is a key requirement and this is associated with age. The proportion in the 45-59 year old category is therefore relatively high. In contrast, in many other occupations the age structures are much more heavily skewed towards younger age groups. In sales occupations, for example, the age structure is much more heavily weighted towards younger age groups. Differences in age structure across occupations influence replacement demand due to occupational mobility and retirement, which are age related. Even inter-occupational mobility is affected differently over occupations;

(b) Retirement.

Retirement rates vary by gender and by age and may differ for different occupational groups. But since sample numbers are often too small to allow for meaningful estimates methods to deal with these problems need to be adopted. Estimates can be based on data from the LFS, which show the percentage of those employed one year ago who have retired from employment, either temporarily or permanently. For males, the main outflows are associated with retirement per se. For females, in particular, there is significant, quite often temporary, outflow for younger age groups associated with family formation;

(c) Mortality.

Another potential outflow is due to mortality. While losses due to death are not great for individual age groups up to the age of 65, they can cumulate to produce significant losses over an extended period. However, the current model does not explicitly incorporate differential mortality risks (not least because no significant or radical changes are expected in them). Rather the focus of the cohort component methodology is to identify overall outflows over cohorts, irrespective of the cause (sickness, death, family obligations);

(d) Migration.

Net migration can have an important effect on the in- or outflow of the labour market. One of the problems of migration is the lack of suitable data;

(e) Occupational mobility.

Occupational mobility is an important source of replacement demand in some occupations although not for all. The full occupational mobility flow matrix indicates that some occupations such as managers tend to gain employment as people are promoted from other occupations. The cohort component approach does not differentiate the replacement demand which is due to occupational mobility. It only identifies net mobility.

The overall scale of change is obviously dependent upon the length of period considered, as well as the opening stocks and the age structure of the current workforce. Replacement demand is also dependent on the level of occupational aggregation. With lower levels of aggregation, the observed occupational mobility is lower. For most projections rates of outflow are assumed to be constant over time. The scale of structural or expansion demand (which in some cases may be negative) is usually modest compared to replacement needs, and in most cases the latter offsets any negative change.

Replacement demand is driven by the proportion of employees in an occupation that is likely to leave within the forecasting period. A higher share of those workers in the base period of a forecast will lead to a higher predicted replacement demand. Business cycles are also prone to shifts and outflows at some point in time. With respect to the timing of (early) retirement, in good economic times people are likely to stay on a little longer; hence the outflow rate will be somewhat lower. In contrast, in bad economic times, people that are likely to leave soon will have incentives to move out earlier.

The period of the economic downturn (global or sectoral) has an impact on the timing of the outflow that can lead to some short-term shifts in replacement demand. Companies try to accommodate the lower demand for workers by reducing flexible work (temporary agents and not core stuff), but also by bringing forward outflows that are likely to occur in the near future (retiring rather than firing). In other words, early retirement schemes – official or not – are being used to reduce the workforce if possible. This implies that the outflow among the cohort of older workers, who are close to retirement, is temporarily higher, while the outflow will be lower for some time after the crisis. This lowering effect is the simple result of the reduction in the population of workers that reach retirement age.

At the other end of the age distribution, the outflow of younger workers is also likely to be increased somewhat as they are more likely to have temporary contracts or may be the workers that are the easiest to lay-off. There can be several interrelated reasons for this, including common ‘last-in, first-out’ rules which apply in many countries or sectors and legal rules which often stipulate compensation based on tenure and age, which encourages employers to focus on younger workers. Other economic arguments of accumulated human capital can also lead to this result.

While the outflow increases temporarily – this in the methodology of replacement demand means that the replacement demand increases as well – it is not likely that replacement demand in those economic circumstances will be filled in immediately. Some catch-up will take place after the economic recovery. The annual replacement rate is taken as an average over the entire forecasting period, which is not affected by such cyclical behaviour on the labour market. However, it is important to be aware of the implications of the crisis for in- and outflows of workers that might lead to a temporary deviation from the overall replacement rate.

## 6.2. Current specifications of the model

The methodology follows the approach used by ROA in its national forecasts (Cörvers et al., 2008), adapted for the data availability of the other European countries. It is based on the cohort-component analysis that uses the EU-LFS for all countries, while disaggregating education into several ISCED categories (for replacement demand by education) and ISCO categories (allowing estimates of replacement demand by occupation).

There are three components to the model:

(a) a forecast of demographic development within a country;
(b) a forecast of (changes in) participation, preferably by gender and age groups;
(c) an estimate of the outflow by occupation (education) category, gender and age group.

Components (a) and (b) are usually considered external to the replacement demand model. Current estimations are based on the baseline model of the Europop 2008 forecast ([28]), which was the most recent demographic forecast by Eurostat. Changes in participation use the same participation rate by country, age and gender as generated within the E3ME model. This insures consistency across the entire set of forecasts.

The basic steps use occupation (subindex o) as the relevant subcategory. However, one can interpret the methodology analogously if education is used instead. For the purposes of the Cedefop forecast, education is not estimated separately given the high level of aggregation on the education variable. Rather, the replacement demand by education is deduced from the occupational replacement demand. By using the occupational replacement demand and

([28]) In the Cedefop forecasting work we intend to use data as current as possible. For the forecast which will be published in 2012, Europop 2010 will be used with the highest probability.

imposing the most recent distribution of education by occupation, we are able to present the most likely replacement demand using the current demand for education levels within an occupational class.

Table 9 gives a schematic input-output table of the labour force/population in a country (see also Willems and de Grip, 1993). The first rectangle gives the movements within the labour market. The second, bigger rectangle encompasses movements out of the labour market, while the third rectangle also considers changes in the population. Adding rows (for time *t*) or columns (for time *t-n*) of these flows gives the total population within an occupation.

Table 9. **Schematic of replacement demand**

Outflows

| **Outflows Inflows** | **Occupation 1** | **Occupation 2** | **Unemployed** | **Outside the labour force** | **Outflow population** | **Total** |
|---|---|---|---|---|---|---|
| Occupation 1 | A | B | C | D | | W1,t-n |
| Occupation 2 | E | | | | | W1,t-n |
| Unemployed | F | | | | | |
| Outside the labour force | G | | | | | |
| Inflow population | | | | | | |
| Total | W1,t | W1,t | | | | |

*Source*: Willems and de Grip (1993).

Several flows are indicated in the table with capital letters. *A* denotes the workers that work in occupation 1 at time *t-n* and continue to do so in period *t*. *B* denote the workers that move from occupation 1 to occupation 2 in the observed time. *E* denotes the opposite movement from 2 to 1. Thus, *B* and *E* denote the job-to-job mobility. *C* and *D* denote movements out of the labour market from holders of occupation 1. Corresponding inflows into occupation 1 are *F* and *G* in the schema.

The first step in modelling future replacement demand per occupational class is a description of the inflow and outflow patterns by occupational class in a historical period. Because there are no appropriate data for mobility flows on the labour market, stock data are used. With the cohort components method, cohort-change rates based on the number of persons of the same birth cohort who were employed at two different time periods can be calculated (Shryock and Siegel, 1980). These cohort-change rates can be rewritten as average annual net inflow or outflow percentages (flow rates for males and females are differentiated):

$$\dot{F}\,^{t-1}_{o,a} = \frac{{}_k\boldsymbol{W}^{t}_{o,a+1} - {}_k\boldsymbol{W}^{t-1}_{O,a}}{{}_k\boldsymbol{W}^{t-1}_{o,a}}$$

(17)

Where $\dot{F}\,^{t-1}_{o,a}$ is the annual net inflow or outflow ratio of workers in occupational class *o* of age group *a* (with class width *k*) at time *t*-1 during the period (*t*-1,*t*); $\boldsymbol{W}^{t}_{O,a}$ is the number of people working in occupational class *o* of age group *a* (with class width *k*) at time *t*. The time lags are used to describe how particular cohort is moving within an occupation. If $\boldsymbol{F}^{t-1}_{o,a} > O$, there is a net inflow for a certain age group from an occupational class, and if $\boldsymbol{F}^{t-1}_{o,a} > O$ there is a net outflow.

The second step in modelling is to translate these inflow- and outflow-percentages into the replacement demand by occupational class. For occupational classes with an increase in employment in the period (*t*-1,*t*), replacement demand is equal to total net outflow in this period. However, for occupational classes which faced a decrease in employment, not all vacancies created by the outflow of workers will have been filled. Therefore, replacement demand for these occupational classes equals the number of vacancies likely to be actually refilled, that is to say the total inflow of workers into the occupational class. In this way, the more or less 'structural' replacement demand is derived. This methodology measures only the net flow to or from an occupational class. This means that replacement demand satisfied by re-entering workers of the same age cohort is not measured. So replacement demand is actually measured only for newcomers to the labour market.

A model is then estimated in which the net inflow or outflow ratios are explained on the basis of the average inflow or outflow from the total working population on the one hand, and the occupation-specific deviations per age-gender group on the other. This approach guarantees that the sum of the net flows among the occupations corresponds to the total inflow or outflow.

Written mathematically:

$$\dot{F}_o = \dot{F} + \sum_k \beta_{ok} D_x$$

(18)

Where:

$\dot{F}_o$ is the vector of net inflow of outflow ratios for occupation *o*, with observations for gender, age group and year; $\dot{F}$ is the vector of net inflow – outflow ratios for the total working population; $D_x$ is the matrix with dummy variables where elements are equal to 1 for cohort $x$ and 0 elsewhere; and $\beta_{ok}$ represents random parameters.

The third step is to project the historically measured net replacement demand rates per age-gender group for a particular occupational class onto the age-gender structure of the workers at the beginning of the forecasting period. The outflow coefficient is then combined with changes in the participation rates and applied to the population of workers within an age cohort. An increase in participation rates implies less replacement demand. Higher participation rates of workers manifest themselves in this model by lower outflow rates. Given that we estimate on historic outflow rates, the estimated coefficient of outflow will be too high for the future. The expected increase in the participation rate is included by correcting the outflow coefficient for these changes in participation rate.

Finally, a projection is made based on the estimated coefficient combined with participation rate changes applied to the age-gender structure of the occupation as predicted by demographic and participation forecasts. To model the demographic composition of an occupation and its dynamic changes, we project uniform changes of one age cohort in the next cohort over the total projected time-horizon.

Replacement demand due to other reasons is the remainder of the estimated replacement demand net of the retirement group. It comprises the occupational mobility, migration, early retirement, and movements out of the labour force for other reasons. This result is therefore driven by all elements of replacement demand, population, participation and the gender-age specific outflow coefficient.

In the forecasting results we distinguish between the two components of replacement demand: retirement and other reasons. While the first figure gives the components by country totals, the second calculates the European totals by occupation group. It can be seen that retirement accounts for less than half of the replacement demand. However, this does not imply that retirement is not one of the most important reasons for replacement demand, as it can also be manifested in some countries as early retirement schemes, or other permanent outflows from the labour force (e.g. unemployment of older workers), which diminishes the population of active older workers before they reach the official retirement age.

An important aspect that influences replacement demand is the participation rate of cohorts by gender. For example, an increase of participation of women around the time that their children reach schooling age (usually in the cohorts of 30-34 and 35-39) will offset other elements of replacement demand. This is a change that has happened or is happening in many EU Member States. Female labour participation is increasing, especially around the age groups that they were formerly retreating from the labour market.

Some Nordic countries, but also the Netherlands, with high replacement demand forecasts have already achieved a high level of participation of women across all age cohorts. They cannot gain as much from increasing participation as can the countries that are starting from a low level of participation for cohorts of middle aged women.

## 6.3. Testing the alternative approaches

The current design of replacement demand modelling in the form of cohort component method is considered as rather robust and stabile. On the other hand the cohort component approach has been criticised in volatile environments, in which for example large inflows or outflows due to migration take place. This critique has been voiced among others by Fox and Comerford (2008).

By using cohort component method we are facing two main difficulties:

(a) data used do not allow us to observe the full population, but rather a sample of the population;

(b) we do not observe birth cohorts, but rather identify them through age cohorts.

Depending on the data set we can distinguish between one year brackets in age, or only five year groups. This last aspect can be important for the derivation of the outflow estimate and it's precision.

### 6.3.1. Panel data

The potential availability of panel data makes it worthwhile to investigate their use in the context of replacement demand. In this section we distinguish between short, rotating panels as they are included in several national labour force surveys, and longer panels that are usually household-based. The theoretical advantage of the panel is that the same person can be observed several times. This allows for a true measure of flows if the population is representative with respect to their occupation change behaviour. Panels are, however, costly to collect and even more difficult to maintain. Such panels, however, might be biased with regard to flow measures, if there is an attrition bias related to flow measures.

The rotating panel is a new element added to many national labour force surveys. It follows up on the LFS with the same respondent (several times) with a subset of questions. Changes in employment and education status can then be used to evaluate the dynamics of those items.

This should allow estimation of the probability of changes in (ISCO) employment and (ISCED) education for a given age and gender group. The short time duration does not, however, allow us to elaborate panel-estimation techniques. It is also potentially more difficult to identify changes, given that changes might also lead to increases in non-response rates. For example, a (respondent) worker migrating to another country will be missing in a follow-up. However, he or she leaves an open position, which we would like to measure within the context of replacement demand. Therefore, a careful evaluation of how missing information within the panels is filled in should be included in the analysis. This might differ by country, resulting in different degrees of accuracy and also outcomes.

The example of the Netherlands is adopted below to investigate how such a rotating panel set-up could be used with the replacement demand.

To examine the usefulness of short panels as an alternative to the cohort component methodology, it is insightful to compare results of flow estimates based on the two different methodologies while using the same type of data. We use in the following the Netherlands LFS which includes a short rotating panel. Each respondent is approached five times, on a quarterly basis. We use the first and last responses which are one year apart. On a 3-digit occupational level, we observe occupational changes.

Any change in occupation is considered an outflow from the occupation at *t* for the occupation at point *t* inflow for the new occupation, observed at *t*+1. Net flows are then the total flows within an age cohort, where we subtract the outflows from the inflows. The percentage flows are then the net flows relative to the total number of workers in the age cohort by gender of an occupation at point *t*.

Table 10. **Regression analysis of cohort component versus panel flow estimates**

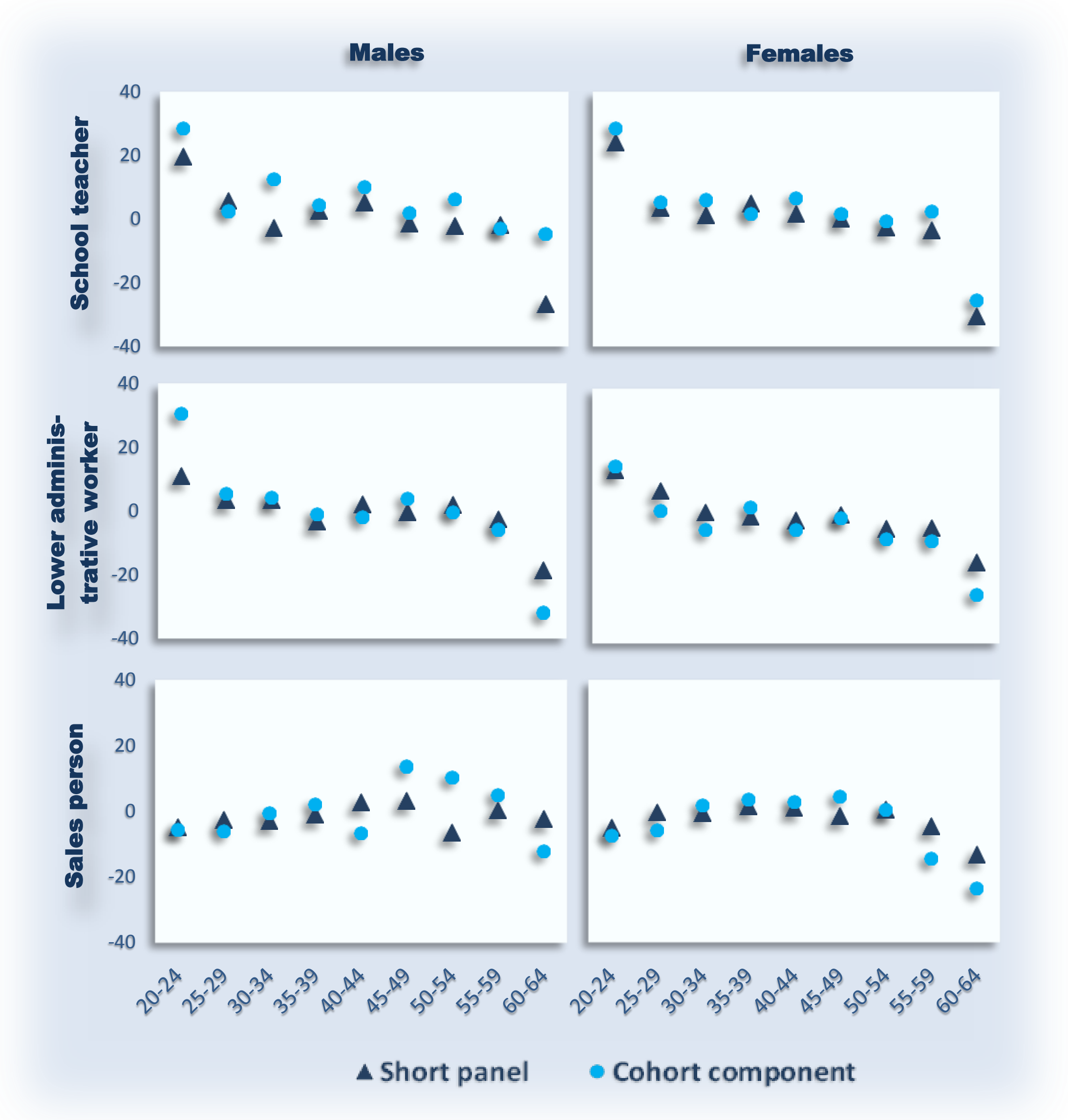
| | % net flow cohort component | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| % net flow panel | 0.667***<br>(0.147) | 0.667***<br>(0.147) | 0.648***<br>(0.143) | 0.858***<br>(0.000316) |
| Gender | | YES | YES | YES |
| Size of occupation | | | YES | |
| Constant | Yes | Yes | Yes | Yes |

NB: Estimates are at 3-digit occupation level by gender and age cohort. Estimation techniques are OLS (1–3) and weighted OLS (4).
*** significant at the 1% level.

*Source*: Dutch LFS, 2001-07, pooled flow measures.

The percentage flows by occupation – age – gender are surprisingly similar. Regressing the net flow rates using panel estimates to those of a cohort component methodology shows a correlation that ranges between 0.648 and 0.858 (cf. Table 10). A similar conclusion can be drawn using the plots of flow rates by age cohort for several occupations. We show three occupations (3-digit occupation level), for the age groups in Figure 7.

Figure 7. **Flow rates for selected occupations, cohort component versus short panel, 2001-07, the Netherlands**




*Source*: Dutch LFS, 2001-07, pooled estimates.

Long panels are usually more accurate as they enable to following up on workers that do not respond ([29]). It is also easier to identify individual labour market behaviour and the changes that appear over its life-cycle and career. However, not all countries have (publicly available) panels which span a longer time period and have sufficient amount of respondents to allow for a more detailed analysis of the behaviour.

Long panels have a small population which is problematic. For each of the occupation categories the outflow by 10 age categories, identified separately for males and females, is identified. This leads to small cell sizes which influence the outflow figures. Given that the population does not rotate as in the rotating (short) panel of the labour force survey, it is also less likely that all occupations are covered consistently. When we attempted to calculate outflow rates for Germany and the UK, it became obvious that the cell size is too small to provide sensible estimates on ISCO 2–digit level. For example, using the last 10 years of the BHPS (1998-2008), and aggregating across gender and ages, yields the outflow percentages by occupation between 0% and 2.9% annually, with an average of 0.26%. Worse still, these flows are on an occupation level heavily influenced by the cell size (the number of people observed in the occupation), in such a way that both the lowest and the highest flow figures are coming from the occupations with the least observed respondents.

### 6.3.2. Retrospective questions

In the retrospective question, workers are asked retrospectively whether the current *t* labour market status is the same as the one they had at some previous time t-1. In detailed questionnaires this can entail changes in employment status, occupation, employer, industry, but also education level and direction.

A problem of this approach is retrospective bias and item non-response. Recall bias is shown to affect non-salient changes in the answers of respondents. Changes that are not seen as being very important are forgotten during the time that the respondent answers the questionnaire. It is likely that the recall bias affects those issues that are salient less than those that are less salient. In other words, movements into or out of employment are probably not forgotten, whereas changes of position, upgrading of degrees, or maybe even changes in companies might be forgotten.

A more subtle problem is the problem that respondents cannot easily judge the timing of events in retrospective questions. If the question is not properly

([29]) As the long-term trends can be used for smoothing eventual non-responses.

anchored, the respondent might attribute different times to changes in his work; however, this recall bias can go either way.

Both issues can lead to a bias in the answer to the retrospective question. Whereas the recall bias can be signed, i.e. giving the direction and the order in which it affects the underlying question, the second bias cannot. We can therefore assume that the errors are random.

### 6.3.3. Administrative data

Administrative data have two advantages: first, they usually encompass the entire population; secondly, they do not rely on surveys that would have to deal with non-response, question wording and other such problems. Therefore, it is usually assumed that administrative data contain little or no error ([30]). Administrative data have been made available to the research community in several countries (cf. Lazear and Shaw, 2009). The data vary, however, in terms of the underlying data source and the purpose for which the data were originally collected. Nevertheless, they do have the possibility to be linked with other data. Several aspects make the general use of administrative data cumbersome:

(a) data sets are generally big;
(b) access to the data is usually very restrictive;
(c) documentation of the data is often limited.

The treatment of administrative data can be done in an analogous way to the long-term panels. The advantage is that there is usually less attrition. However, given that access to administrative data is restrictive, and often only available within national borders, we refrained from continuing and including exemplary flow rates.

### 6.3.4. Conclusions from testing the alternatives

The alternative approaches and data sets described above can be used in forecasts at national level but their usability for Cedefop forecast is limited due to lack of necessary data for the whole of Europe.

There could be a way if the EU-LFS would systematically ask the retrospective question in all countries. Given that a similar question on the sector

([30]) The degree to which there are no or few errors in the data depends, however, on the purpose for which the administrative data are collected. For example, the data of the Netherlands are based on Dutch tax data. They contain only data on legal salary payments, including all errors made by companies in calculating them. Given that the tax authorities have every incentive to control such errors, it is likely that the data are at least more accurate that those contained in surveys. They would be, however, less accurate for the purpose of assessing money paid for (illegal) work.

of employment in the previous year exists, it seems rather straightforward to include a similar question on the occupation. Ideally, also the practice of including short, rotating panel set-ups for the national LFS could be applied for the whole of Europe. As the case of the Netherlands show, this would provide a more stable estimate of net flows and also allow further investigation of the causes of the flows.

Another option would be to allow national experts to supply unconventional flow data using one of the alternative methodologies that can improve the flow measure. Good and consistent national data sets could potentially significantly improve the quality of the replacement demand forecast for Europe.

CHAPTER 7.
# Reconciling skills supply and demand

Economic theory as well as economic reality shows that coexistence of unemployment and unfilled vacancies is a natural and simultaneous feature of the labour market. Even in full employment it is difficult to reduce the number of unfilled vacancies to zero due to the search costs and asymmetries in the labour market. This is especially the case when the imbalances within particular occupations are caused by lack of appropriate skilled labour force, which is usually solved by retraining and/or reallocating workers. One of the aims of the process of anticipating changing skill demand and supply is to uncover potential imbalances in such a way that adequate preventive measures can be taken.

Knowing more about potential skill imbalances in the labour market is very important for policy. The projections of the skills supply and demand can throw light on possible future developments in European labour markets, highlighting potential education mismatches and helping to inform decisions on investment in skills, especially formal qualifications, made by individuals, organisations and policy-makers.

Comparing current demand and supply projections is, however, problematic for both practical and theoretical reasons. Unless the two sets of results are based on common data and are carried out simultaneously, they cannot be directly compared. In reality there are many adjustment mechanisms that operate in the labour market to reconcile imbalances that may arise. In the short term, these include adjustments in wages and different kinds of mobility, as well as changes in the ways employers use the skills that are available. In the longer term both supply and demand will adjust to reflect the signals and incentives that arise due to these forces.

Generally, employers will not cease their operations if they cannot find the ideal mix of skills. They will operate with what is available. Conversely, if the educational system delivers too many people with particular levels of formal qualification this does not necessarily imply that these people will remain unemployed. Rather they tend to find jobs that make less direct use of their qualifications, since their education and training still often puts them at an advantage in the labour market compared to those with lower level qualifications. The labour market operates as a kind of sorting mechanism that allocates people to jobs, based on the limited information available to both sides (employers as well as actual and potential employees).

In principle, these mechanisms can be modelled and incorporated into a forecasting tool. In practice, this demands very detailed and rich data that are, at present, not available at pan-European level. In the present results, a more limited reconciliation is used to recognise the key features of the labour market interaction. The projected numbers on the supply side are taken as given for this purpose. This reflects the fact that the total numbers available by qualification level are largely predetermined by demography and educational and training decisions already made. The much better qualified new entrants coming into the labour market, replacing much less well qualified older people, imply that substantial improvements in average qualification levels are inevitable in the short to medium term.

This chapter describes Cedefop solutions to the above-mentioned problem. The main data and methodological limitations of reconciling labour supply and demand with respect to qualifications, including the use of labour market accounts residuals are presented in Section 7.1. Section 7.2 focuses on the approach used by Cedefop to provide reconciliation of skills supply and demand. The construction of possible indicators for better understanding and interpretation of the imbalances is described in Section 7.3. More details can be found in Kriechel and Wilson (2010, 2011b).

## 7.1. Measuring imbalances

Within the current Cedefop framework, the focus is on employment in sectors, occupations and by qualifications; labour supply by qualification, age and gender; and unemployment distinguished by qualification category. Vacancies are not used as there are no detailed and comprehensive ([31]) measures available at pan-European level ([32]).

When comparing estimates of demand and supply of labour (and skills) it is important to recognise significant problems that arise due to different measurements applied in different sources. This section focuses on differences in the historical estimates of the labour force, employment and unemployment used in the Cedefop framework.

([31]) The vacancy monitor used by the Directorate-General for Employment of the European Commission does not provide data of sufficient quality to be used in the modelling framework.

([32]) Although Cedefop is exploring the possibility of collecting such data via a new employer survey.

The population data used in E3ME are constrained to match the official Eurostat numbers and projections. The labour supply numbers in E3ME are therefore not LFS benchmarked as such, although they use LFS activity rates. They rely on Eurostat demographic data to produce the overall numbers and LFS activity rates by age and gender.

Labour demand is measured as employment. Obviously this is not strictly correct. Observed employment is the consequence of both demand and supply factors, nevertheless it is common practice in work of this kind to refer to employment as de facto 'demand'. Overall employment levels from E3ME by sector are translated into implications for occupations by the EDMOD module, as described in Chapter 5. This is based on a detailed analysis of occupational employment patterns within sectors. This module also produces initial (unconstrained) estimates of employment by qualification (again based on an analysis of changes in employment patterns within occupation and sectors).

To focus on imbalances by skill (as measured by qualification), these initial estimates of employment by qualifications are constrained to match 'supply in employment'. This is a measure of the number of people by highest qualification held who are economically active and in employment. This is on an EU-LFS demographic accounts basis. It is then scaled to match the E3ME national accounts (NA) based estimate of employment. This final set of employment estimates by qualification matches the total employment in E3ME on a NA basis and is known as 'constrained demand'.

The LMAR in E3ME is the difference between employment (workplace jobs, NA) *E(j)*, and employment (head counts, LFS based, residents) *E(r)*.

$$LMAR = E(j) - E(r) \tag{19}$$

The person/jobs distinction is not entirely clear cut. The LMAR arises for a range of reasons of which occupation is just one (double jobbing). Other factors include: commuting (flows across national borders) and statistical discrepancies between NA and LFS based estimates. NA employment appears to generally refer to persons rather than jobs (as in the LFS), meaning that most of the discrepancies arise from reasons other than double jobbing.

The LFS based estimate of employment (*E(r)*) is a count of people employed in a particular country. It is equal to the labour force (LF) less the number of people unemployed (U), both measured by head count.

$$E(r) = FL - U \tag{20}$$

Information on unemployment by qualification level is available from the LFS. This can be used in combination with the projected totals of unemployment from E3ME to generate projected levels of unemployment by qualification level, by making assumptions about maintenance of historical differentials in unemployment rates.

Equations 18 and 19 can be variously rearranged to derive indicators of interest from the model outputs:

$$E(j) = LMAR + (LF - U) \quad (21)$$

In E3ME the following slightly different terminology is used:

$$E(j) = EMP \quad (22)$$

'Unemployed' in E3ME is given by:

$$Unemployed = LF - E(j) = E(r) + U - (LMAR + E(r)) = U - LMAR \quad (23)$$

E3ME then generates a variable called 'U level' as follows:

$$U\ level = Unemployed + LMAR \quad (24)$$

This is equivalent to the normal ILO measure of unemployment.

Unemployment is calculated this way in all countries, except for Luxembourg where the discrepancies were so large that it was not possible to model unemployment in a stable manner. The LMAR is calculated to match the last year of historical data, taken from AMECO database of European Commission's Directorate-General for Economic and Financial Affairs, which uses ILO-consistent definitions. In the modelling, the LMAR is calculated using a 'fixed share of employment', although in the workbooks it can appear to vary; this is because the unemployment rate in the workbooks is only calculated to one decimal place. 'U level' can easily be calculated from 'U rate' by multiplying by the total labour force (LF).

However, as shown in equation 22, this will include the LMAR. The value of the LMAR will vary for those with high, medium and low qualifications. It is possible to assess this by going back to the raw LFS data and using the historical estimate of the unemployment rate for high-, medium- and low-qualified people to generate a number for those unemployed in each category in the base year. This can then be compared with the difference between the demand and supply numbers in the base year. The difference is the LMAR for that qualification category. In principle, this can then be projected assuming (in the first instance) that it remains a constant proportion of the total LMAR (from E3ME) for all future years. In practice this has not been done. Instead an assumption is made about how unemployment is 'shared out' among high, medium and low qualification categories.

The E3ME data on total unemployment are used to provide overall constraints on the implied unemployment numbers by qualification (three broad levels: high (H), medium or intermediate (M), and low (or no) qualifications (L)) in the imbalances work. The latter are constrained to match the overall E3ME estimates of 'U level', making assumptions about how overall unemployment will be 'shared' between those with different levels of qualification, using LFS historical data and assumptions about how these patterns might change in the future. These assumptions are that previous differentials will be maintained but that the better qualified will take an increasing share of total unemployment in line with their increasing share of the workforce.

One measure of initial imbalance and mismatch by H, M and L qualification levels can be obtained by subtracting the unconstrained 'demand' estimates (employment estimates by qualification from the demand model, summed across all sectors and occupations) from the corresponding 'supply' estimates (the total labour force across all ages and genders).

Initial comparisons suggest that the independent projections of qualifications patterns for supply and demand exhibit significantly different trends, with the supply of those with high and intermediate level qualifications generally rising more rapidly than the demand.

The implied demand side employment rates by qualification category, when compared with the supply side activity rates, indicate some odd patterns, both across categories and over time. In principle, of course $e/r < a/r$. A final adjustment has been made to the employment by qualification estimates to take into account the LMAR.

## 7.2. Estimation of imbalances

The historical patterns of employment by qualification observed are the result of a combination of both supply and demand factors. Separating them is not straightforward. Recent trends have seen a sharp rise in the formal qualifications held by those in employment in most countries. There is some evidence that this reflects demand changes, with many jobs requiring more formal higher level qualifications than used to be the case. There are also indications that the returns on such qualifications have remained high (for a review, see Wilson et al., 2007). On the other hand it is clear that there have been major changes on the supply side, in part at least in response to government policies to increase participation in higher education. The latter has resulted in a big increase in the numbers of people entering the labour market with high formal qualifications. The proportion of young people with higher formal qualifications is much higher than for older people. There is, therefore, a strong cohort effect. This has been reinforced to some extent by increasing qualification rates for older people as well (due to lifelong learning and upskilling effect) (part of Section 3.4).

The observed patterns of employment (stocks of people in employment with formal qualifications) clearly reflect both demand and supply side influences. Certain indicators are more informative about one than the other. In particular, there are various measures of the flows of people through the education system which can be regarded as primarily supply side indicators (although even these reflect decisions that people are making about education based on their perceptions of the overall balance of supply and demand for different qualifications).

In principle, it is possible to develop quite sophisticated analyses of the numbers of qualified people at higher level, using information on the flows of people through the education system. The overall supply of people holding formal qualifications at higher level (ISCED 5-6) is relatively straightforward to conceptualise and model. However, there are considerable conceptual and practical difficulties in extending this to include lower-level qualifications (ISCED 1-2). A particular difficulty is to conceptualise the idea of supply to cover specific dimensions such as occupation, sector and geographical area. This is because the educational systems in most countries are not completely hierarchical. These issues are discussed in turn:

(a) limitations by ISCED level:

the first problem to be addressed in extending this type of model to cover the full range of qualifications is the much more limited information available on lower level qualifications (ISCED1-2). Ideally, stock-flow modelling requires a comprehensive set of demographic accounts showing how individuals progress throughout the educational system and the labour market over time. In practice such accounts do not exist, although there is a considerable amount of information on certain flows as exploited in the skills supply module described in Chapter 4.

Statistical agencies collect and publish a considerable amount of information on the higher education in particular. This can be used to develop estimates of the main flows involved and thereby develop stock-flow models of this process. In practice, this information is often disparate and far from comprehensive.

In the case of lower-level qualifications, while there is an enormous amount of detailed information available on the acquisition of qualifications, there is much less information on what prior qualifications these individuals may have possessed. This makes it difficult, if not impossible, to develop stock-flow models analogous to those constructed for higher levels (e.g. for the UK in Wilson and Bosworth, 2006).

(b) highest versus all qualifications held:
the discussion so far has focused on highest qualifications held. As individuals acquire ISCED 4, 5 and 6 qualifications, it is almost inevitable that the proportions with ISCED 1-3 as their highest qualification will fall. This can mean that, despite increases in those acquiring ISCED 1-3 qualifications, the numbers and proportions of people possessing these as their highest qualification may actually decline.

(c) problems in conceptualising supply into occupations or sectors:
most occupations are undertaken by people with a range of formal qualifications. This is partly a function of age, with older workers generally relying more upon experience than formal qualifications. However, even allowing for the age factor, there are enormous differences. This makes defining the supply of people into an occupation almost impossible. It is possible to identify some key elements, focusing on the flows of people through the education and training system, but boundaries are too blurred and transitory to enable robust quantitative modelling.

Much the same is true for the concept of the supply of labour to a sector. This will depend upon the occupational mix of the sector and its geographical location. For some occupations the labour market may be worldwide. This is increasingly true of many high level managerial and professional groups. Ever increasing ease of transport now means that it is also a feature of the labour markets for many lower-level occupations (for example, construction and agricultural workers, as well as nurses). While these issues may be addressed within individual sectors, it is very difficult to develop a general approach that can cover all these aspects consistently for the whole economy.

There may be some overqualification or underqualification and this pressure varies among occupations and sectors. The following elements are important here:

(a) the demand for qualifications model, which delivers overall numbers of people in employment, qualified at three broad ISCED levels;
(b) the stock of labour supply;
(c) the sorting model (SORT algorithm described in the Annex 4), which sorts people according to the qualifications held into occupations to make the results from (a) and (b) consistent.

The SORT algorithm reconciles the two sets of estimates of demand for and supply of qualifications. This final element compares the supply numbers with the demand ones and re-computes the employment patterns to bring the two into agreement (making certain assumptions about unemployment). Effectively it acts as a sorting mechanism that raises or lowers qualification shares within occupations until demand and supply numbers match. This does not imply that demand and supply are equal however, since some people may be overqualified or underqualified for the jobs they are employed to do.

The Sort algorithm used in BALMOD ([33]) provides a simpler approach to reconciling the demand and supply results in aggregate, given the data available for all countries. Its outcome is that for individual occupations or sectors the patterns of qualifications as revealed by the original unconstrained demand projections and the constrained ones will show how any surpluses or shortages affect the qualification mix. If supply is growing faster than demand for particular levels of qualifications, the constrained qualification mix will be ‘richer’ than the unconstrained ([34]) one (and vice versa). Comparison of the constrained and unconstrained results provides a useful indicator of supply-demand pressures for different occupational and sectoral groups.

The concept of the supply of qualifications at spatial level is somewhat more manageable than it is for occupations or sectors. It is relatively straightforward to develop quantitative estimates and projections of population and the labour force for each country. In principle, this can be extended to cover formal qualifications held. However, the data available at pan-European level are generally less robust than at national level. Moreover the issues of commuting and migration flows become significant. While it is possible, in principle, to envisage the development of customised qualification supply models for each individual country this would require considerable resources and time.

The present modelling is, therefore, limited to a much more simplified level than the more detailed and sophisticated stock-flow analysis applied in some individual Member States (e.g. the UK, Wilson and Bosworth, 2006).

## 7.3. Indicators of imbalances

In addition to the reconciled demand and supply estimates, indicators of imbalances were also developed to help to understand how imbalances evolve,

([33]) Module BALMOD of the Cedefop conceptual modelling framework (Figure 1) is designed to bring together the demand and supply estimates to tackle this issue. Demand is projected by sector, occupation and qualification level (highest qualification held), whereas the supply focuses on qualification by gender and age.

([34]) As unconstrained demand we understand demand initially produced by the model. As constrained demand we understand demand adjusted to fit the supply.

to identify emerging trends and to allow for comparison among countries. These indexes provide a comprehensive picture of possible imbalances and mismatches in the labour market. We distinguish between indicators that are derived from the raw supply-demand comparisons and those that are derived after we have reconciled supply and demand. These indicators are currently in the stage of preliminary results and their robustness is being tested.

#### 7.3.1. Unemployment as a general indicator

The first set of indicators is based on the supply and demand forecast from the separate modules. These indicators are indicative of overall imbalances in supply and demand, i.e. an oversupply or undersupply, be it by education or in total. Furthermore, we use the extrapolated unemployment rates by education level, both as an indicator of the allocative process but also of imbalances that are not directly identified within the model.

The main relation of labour supply and demand is that supply should be equal to the sum of fulfilled demand and unemployment. As described in previous sections, unemployment is derived within the model. The unemployment rate by education level can be seen as an indicator of matching efficiency. Overall high levels of unemployment imply that the supply cannot be properly matched to demand. This can have two reasons: (a) there is too much supply relative to demand; (b) the supply does not meet the requirements of the demand. The first point is straightforward as it a simple imbalance of total oversupply. The second one captures elements of the efficiency of the labour market in matching and generating the supply of skills that the economy actually needs. For example, there can be a high overall unemployment rate if the labour market demands only highly educated workers, while the supply consists mainly of people with low and intermediate levels of education. If firms do not adjust their production technology to substitute low- and intermediate-educated individuals for highly educated ones, in other words change their demand, it will lead to unemployment among the low and intermediate educated, while at the same time shortages there will be of highly educated workers. Within education levels there can also be mismatch if the field of education within a level does not match supply and demand. For example, if the supply is mainly within the field of health, while the demand is mainly in the field of technical studies, it is likely that there will be at least temporary unemployment. This last aspect of field mismatch is not currently modelled in the supply and demand forecasts. If, however, this is a structural problem within the economy, this will show up in higher unemployment (and vacancy) rates.

### 7.3.2. Difficulty to hire indicator

This indicator should identify the difficulties an organisation is likely to have if it needs to hire a worker for a specific occupation. The indicator (ITKB) summarises the supply-demand relationships of all education types that can be employed for each level of education. This is weighted by the likelihood that an occupation is filled with a certain type of education. The weighting is based on observed (base year) shares of the occupation-education matrix. The number of people working with a specific background (education) in an occupation thus determines these weights.

$$p_i = max\left(1, \frac{demand}{supply}\right)$$

$$ITKB = \frac{\sum_i p_i x_{ij,t-1}}{\sum_i x_{ij,t-1}} \quad 0 \le ITKB \le 1 \qquad (25)$$

Where $x_{ij,t-1}$ is the total amount of people in occupation *j* with education type *i* in the base year *t*-1. The indicator gives the relative degree to which difficulties can be expected in meeting the occupation in demand. Note that the share $p_i$ is the same for all occupations, as it simply denotes the relative demand to supply of an education type. The implicit assumption is that all shortages in education types will be felt in the same way by all occupations, but weighted to the importance of that education type for the respective occupation. A value of 1 indicates that there are no shortages expected, whereas a (theoretical) 0 would indicate that no demand could be fulfilled. The indicator is rank-ordered and the quintiles determine the relative level of difficulty to fill a vacancy in that occupation.

### 7.3.3. Occupation level indicators based on RAS outcomes

*Indicator of change (IC)*

The indicator of change measures the adjustment that is needed to unconstrained demand (*Du*) in order to reach the level of constraint demand (*Dc*). High levels of the indicator of change indicate significant adjustment processes necessary and higher potential for imbalance. The indicator is calculated for each occupation *j* across all education types *i*.

$$IC_j = \frac{\sum_i |D_{c,i} - D_{u,i}|}{\sum_i D_{c,i}} \qquad (25)$$

The indicator shows the level of change that is necessary relative to the path that the occupation is taking in the projection years. Higher values indicate a higher level of constraint and more adjustment to the current path of employment.

*Measure of change (MC)*

A similar indicator is the measure of change. It calculates the distance between constraint demand (*Dc*) at the year of the forecast (e.g. 2020) and the base year counts (*D*1) (e.g. 2010). In other words, it gives the adjustment that is necessary from base year to the forecast.

$$MC_j = \frac{\sum_i |D_{c,i} - D_{1,i}|}{\sum_i D_{1,i}} \tag{27}$$

Higher values indicate higher level of adjustment relative to the current (base year) state of the labour market and indicate higher potential for imbalance.

CHAPTER 8.
# Quantitative evaluation of the skills supply and demand forecast

The main motivation for evaluating Cedefop's forecast was to examine its consistency, robustness and usability. The rationale for such work has been hotly debated from the very beginning, with many critics arguing that such exercises are neither useful nor necessary. Such criticisms have been rejected by those involved (and others). Some of these rebuttals are technical, but many arguments are along the lines that such critics are missing the point and often end up attacking a 'straw man'. The aim of this chapter is not to directly extend that particular debate, although it is concerned with certain key issues that have been fought over by employment forecasters and their critics. These include the question of how to evaluate better employment forecasts.

In some respect the evaluation could be regarded as premature. The main work has only just begun and insufficient time has elapsed to properly assess projections which still focus upon the period yet to come. However, it is important to put in place systems to assess the work and to learn lessons from this in order to improve future projections. The process of evaluation is not solely a quantitative technocratic one, but also entails a more qualitative element focused on the use that is being made of the projections and their value (or otherwise) to users. The approach set out here attempts to encompass both quantitative and qualitative assessments, from both an internal and external perspective (the latter drawing heavily upon the Skillsnet network of individual country experts (ICEs), who have provided critical support from the start of the process).

Before focusing upon the specific quantitative and qualitative assessments of the results produced to date it is important to deal with some general issues on what it means to talk about forecast accuracy in the context of the social sciences. It is argued that such a concept may be a chimera and that the focus of attention should not be on some non-operational ideal measure of forecasting precision. Instead, a more general emphasis should be given on the usefulness and value of the projections.

As becomes clear from the discussion below, a key aspect of the quantitative evaluation is the accuracy of the macroeconomic model, which underlies the employment projections (as is the case for nearly all exercises of this nature worldwide). This chapter is based on background paper of Wilson and Pollit (2011). The first part of this chapter is focused on the general ideas underlying the labour market forecast evaluation. The second part describes the focus and objectives of this exercise. The third section will discuss the overview of projection errors. The final part presents results of the quantitative evaluation. The results of particular indicators (economic development, supply and demand) will also be presented.

## 8.1. Considerations in evaluating labour market forecasts

Labour market forecasting should be regarded as part of a continuous process rather than a one-off exercise. As such, the projections and the models used to produce them are continually being assessed and modified in the light of new information. Nevertheless, it is useful for users of the results to have some idea of the accuracy, reliability and robustness of the forecasts.

Assessing the accuracy of relatively simple equations and systems is a well-established branch of econometrics. Such analysis is often undertaken as a partial contribution to evaluating the overall accuracy of sets of employment projections. However, it is quite a large step from a straightforward academic exercise to assessing the accuracy of more complex systems which typically lie at the heart of most detailed employment projections. These are generally dependent upon models involving literally thousands of independent (often interrelated) equations.

The most obvious method for assessing the accuracy or reliability of a forecasting system is to consider how well forecasts made in the past matched up to reality. However this fails to recognise two key points. First, a fundamental problem in all social science forecasting is that forecasts are generally intended to influence behaviour and change outcomes from what might otherwise have been expected. The second one is based on an almost philosophical question about what is meant by ‘reality’.

The first point means that evaluation of accuracy is not a simple task. The situation is not one where the results of a simple scientific experiment can be accurately measured, given fixed and known conditions. Such considerations call into question the validity of asking about how accurate the projections are at all. A more relevant question to ask is whether or not the forecast was useful.

This can be illustrated by a very simple example. Consider a forecaster, taking a stroll by the seaside, who observes a blind person walking towards a cliff edge. The forecaster might conclude that, on the basis of observed behaviour, assuming the blind person does not change this pattern of behaviour, then he or she will fall over the edge. If the forecaster shares this forecast with the blind person, who as a consequence changes direction and avoids disaster, the forecast would turn out to be 100% inaccurate. However, it would not have been useless. Indeed, assuming that the blind person was not bent on suicide, it would have been very valuable.

Of course, the consequences of most social science forecasting are not a matter of life and death. Nevertheless, they are frequently concerned with highlighting possible problems and alerting policy-makers and individual labour market participants to the consequences of current patterns of behaviour and policy.

The second point raises concerns about data quality and what ‘reality’ means. All too often in the social sciences analysts are faced with the problem of working on shifting sands. Official estimates of fundamental economic and labour market indicators can often change very dramatically from one vintage to the next, changing perceptions of historical reality. Such shifts in understanding the current position and what recent trends have been can clearly make the task of predicting the future very uncertain. It is essential that such factors are considered when trying to assess the accuracy and reliability of any projections. This is a particular concern in the pan-European projections which are based on less sound statistical foundations than many national ones.

## 8.2. Focus and objectives of the evaluation

The models used for developing occupational projections are generally very large and complex, often appearing to the uninitiated as black boxes which produce results which are not easily traced back to small set of simple assumptions or axioms. One of the aims of the present study is to increase the accessibility and transparency of the employment projections produced as part of the Cedefop work. A second aim is to provide some useful information about the value, reliability and accuracy of such projections.

The demand for labour depends on many factors, including the state of the macroeconomy, the scale of capital investment and its allocation among industries, the pace of technical change, changing patterns of tastes as real incomes rise, and government policy. These factors are often interconnected. Developments in one industry (e.g. the introduction of IT in the service sector) affect the demand for labour in other industries (in this case, by raising the demand for the output of those industries that produce computers and related equipment). The multisectoral macroeconomic models used by most employment forecasters to produce their projections attempt to incorporate all these factors in a set of consistent and interlinked forecasts for labour demand, covering the whole economy.

The main focus of the present evaluation was on the occupational employment dimension. However, to address this issue it was also necessary to consider the various factors which drive occupational projections. These include the multisectoral macroeconomic model outcomes and, in particular, the industrial employment forecasts which drive the occupational employment projections. In each case, the outcome may be different from the projections for a number of reasons. It is important, therefore, to have a clear understanding of how the occupational projections have been derived.

The main and most obvious aim of this evaluation was to assess the scale of forecast errors in the key variables of interest. A secondary and equally important objective is to attempt to explain these, so as to inform users of the projections about the sensitivities of the results to different factors and to give them some feel for the range of possible future forecast errors.

However, given the present state of economic knowledge, it is clear that any attempt at disaggregated medium-term economic forecasting will result in mixed success. Whatever measure of accuracy is chosen, the forecasts for some variables are likely to turn out to be more accurate than others. Forecasters should be wary therefore of implying undue precision to their results. Rather they should emphasise their value as an attempt to work out the possible implications of a set of explicit assumptions about the future which users may find helpful.

## 8.3. Overview of projection errors

Given the way employment projections are usually produced and the concepts and issues raised above, several factors contribute to the errors observed in occupational projections. The ideal taxonomy of these errors is in Annex 5.

The cause of error is that primary data used at the time of the forecast are likely to have been erroneous and may have been revised subsequently; such errors will clearly affect the forecasts. Error in the measurement of the dependent variable is obviously crucial. There are many examples of major revisions to employment estimates. These affect all dimensions (industry, gender and occupation). Errors in the endogenous drivers of the dependent variables are likely. For example, in the case of employment, variables such as economic output or wages will affect the projection of that variable. Errors in the assumptions about exogenous variables underlying the forecast may have been incorrect too. These may be technical assumptions such as extrapolated shares in the occupational model or policy assumptions in the macroeconomic model.

Errors may also arise because of the forecaster's errors. Errors of judgement by the forecasters about exogenous variables; errors of specification relating to various aspects of model specification (the models may themselves be mis-specified and inaccurately represent the way that the economy and labour market behave). This is reflected in the various parameters of the model. In addition, there may be further errors of judgement made by the forecasters in modifying the model's output. Econometric and other forecasting models are rarely used without some intervention by the operator to deal with perceived problems.

The various factors set out above are often interrelated. Errors in data may cause model misspecification or the need for operator judgement. Nevertheless the above taxonomy provides a means of explaining how observed outcomes may differ from projected ones.

Because the occupational analysis is towards the end of a sequence of steps, involving development of a macroeconomic scenario and sectoral prospects, the occupational projections will depend on the accuracy of a whole series of prior models and assumptions. These are represented in Table 11. They can be divided into three main categories:

(a) macroeconomic/scale effects concerned with the overall level of employment;
(b) sectoral effects, dependent on the factors influencing the broad industrial employment structure of the economy;
(c) occupational effects, primarily driven by technological and organisational change.

Although the main focus of attention here is the bottom right hand corner of the table (total errors in the occupational projections), this depends directly on all the cells to the left and above, any one of which may contribute to the final error. Generally, the overall error in the bottom right hand cell could be broken down into component parts from all the other cells. The main interest is not in the causes of the errors in the rows above, but their total scale. Therefore the focus here is on the right-hand column and the bottom row only. In other words, the concern is not so much to determine the possible source of the macroeconomic or sectoral structure errors, but mainly to identify their average level.

Various summary statistics can be generated, which throw light on the accuracy of the projections. A very simple indicator of accuracy is a comparison of the predicted and observed direction of change. This can be produced for both industries and occupations. Some care in interpretation is needed in cases where little change is occurring and so the results may simply reflect sampling errors in the data.

Table 11. **An ideal taxonomy of the main components of error in developing occupational projections**

| Variables | Data errors: (basic sources (indicators), classification changes: NACE, ISCO, ISCED, etc.) | | | Forecaster's errors | | Total error |
|---|---|---|---|---|---|---|
| | Errors in data on dependent variables | Errors in data on endogenous variables | Errors in data on exogenous assumptions | Judgements on exogenous assumptions | Model error (parameters) | (*ex post*) |
| **1. Macro level/scale effects** | | | | | | |
| 1a. Macroeconomic | GDP | Prices, wages, components of aggregate demand | Public finances, world economy exchange rate demography | Changes in key indicators | Macro model parameters | GDP growth Productivity growth |
| 1b. Total employment | Total employment | GDP | Aggregate labour productivity | Productivity adjustments | Macro model parameters | Total employment growth |
| **2. Industry, etc.** | | | | | | |
| 2a. Detailed employment by sector | Sectoral employment | Output, wages | Sectoral productivity | Sectoral productivity trends | Econometric estimates | Sectoral employment growth |
| 2b. Employment by type (gender, status) | Employment by type | N/a | Employment by sector | Trends by type | Coefficients (shares) | Shares by type and gender |
| **3. Occupation** | | | | | | |
| 3a. Detailed employment by occupation | Occupational employment | N/a | Employment by sector and by type | Trends by occupation | Coefficients (shares) | Shares by occupation |

Another useful summary statistic is the absolute (i.e. discounting the sign) error in percentage per annum growth rates (ppagr) ([35]). It is the difference between the observed ppagr and the projected one. To place this figure in context, it can be compared to the observed typical values for ppagr (again ignoring the sign). However, this ratio can take extreme values when the actual change is close to zero.

Shares of employment measures the projected share compared with the observed share. A particular issue here is that historical revisions can affect the starting point quite significantly. It is also helpful, therefore, to consider a second measure, which nets out any errors in the base year share due to later data revisions.

Finally, it is possible to focus on errors in the projected level of employment. The absolute (i.e. discounting sign) percentage errors can be computed by comparing the projected and observed levels and expressing these as percentages (ignoring sign differences). Again, corrections to the base year data need to be considered.

The errors in the occupational projections are linked to those at industry level. The complexity of this relationship makes it hard to construct simple descriptive statistics, since this requires detailed information about the occupational mix within industries. Nevertheless, some patterns may emerge. For example, faster than average jobs losses in sectors such as mining and manufacturing may be an important factor in the underestimation of job losses among certain skilled and other manual occupations (given their high concentration in these sectors). Similarly the underestimation of increases in employment in some service industries can be a major factor in the underestimation of job increases among particular occupations (e.g. managers or caring personal service occupations).

## 8.4. Evaluation results

The analysis has been carried out by reviewing the forecasts published in the first two pilot projects, Cedefop (2008) and Cedefop (2009), and comparing these to actual outcomes and to current projections (which are based on more recent data).

([35]) It is important to focus on absolute errors since otherwise the pluses and minuses tend to cancel out margins of error for aggregates are, as a consequence generally less than those for more disaggregated categories.

The primary evaluation analyses were focused on Europe as a whole, rather than individual countries, although there is some discussion of the situation in the larger Member States. When the original demand-side projections were formed in 2007, Bulgaria and Romania were not yet covered by Cedefop forecasts ([36]), and so the EU-25 is used as an EU total and EU-10 as a total for the Member States that joined the EU in 2004.

The projections of labour demand (employment) come from external projections of economic output (by sector) and productivity. The performance of the output forecasts, therefore, has a crucial bearing on the performance of the employment forecasts. First, comparisons are made between the projections for growth in economic output and the actual outcomes. This is followed by a brief section on the other explanatory factors for employment demand and a comparison of output and employment projections by sector.

#### 8.4.1. Output

The original source for future output growth used in the 2007 baseline projections was the *European energy and transport trends to 2030* (2007 update) (European Commission, 2007) report, which the current version of E3ME uses as its baseline. The values shown in subsequent tables (especially for the 2006-08 period) are produced by Cambridge Econometrics, but using the European Commission publication as a starting point and are broadly consistent at macro level.

Figures in Table 14 clearly show that the 2007 projections overestimated growth for the period 2006-08. For example, growth for the EU-25 was forecast to average 2.7% p.a. but growth was actually 0.8 percentage point lower, at 1.9%. The relatively low rate of growth over the period is largely attributable to the effects of the financial and economic crisis and falling growth rates in the second half of 2008, which were not anticipated in the 2007-based baseline ([37]).

Similarly, the 2007-based forecasts for output growth over 2006-15 were also much stronger than the current projections. Although the long-term post-recession trend rate of growth is largely unchanged, average annual growth over the period after the recession in 2009 (2010-15) is around 2.1%. The revised figures include reduced output in 2008-09, lowering the average growth over the period. Output growth is currently expected to be around one percentage point

([36]) The focus was to examine results coming out from E3ME model (Chapter 3).

([37]) Romania and Bulgaria have joined the EU in 2006 and so due to temporary lack of consistency and unavailability of data they were included in Cedefop's forecast in 2009. Results were published in Cedefop (2010).

lower each year over the period than in the 2007-based forecasts, with Spain being the worst affected country.

Table 12. **Predicted output growth in 2007 compared to actual outcomes and current projections**

| | Average annual gross output growth, 2006-08 | | Average annual gross output growth, 2006-15 | |
|---|---|---|---|---|
| | Expected then | Actual** | Expected then | Expected now |
| EU-25* | 2.7 | 1.9 | 2.7 | 1.7 |
| EU-15 | 2.5 | 1.8 | 2.6 | 1.6 |
| EU-10* | 5.2 | 5.1 | 5.0 | 3.8 |
| DE | 2.7 | 2.2 | 2.4 | 1.1 |
| FR | 2.7 | 1.5 | 2.8 | 1.7 |
| UK | 2.5 | 1.6 | 2.6 | 2.0 |
| IT | 1.2 | 0.4 | 1.7 | 1.4 |
| ES | 3.5 | 2.5 | 4.2 | 2.0 |

* Bulgaria and Romania were yet to be included in E3ME at this stage.

** The measure of output is gross, as this is the determining factor in E3ME's equations. However, data availability is better for net output (GVA) in recent years so this is used for 2006-08 actual data. The difference between the two measures in terms of growth rates is small.

*Source*: Wilson and Pollitt (2011).

### 8.4.2. Value added

This analysis focuses on the forecasts for 16 broad sectors (built up from more disaggregated sectoral forecasts). The sectoral breakdown is important for the analysis as the model's projections for aggregate employment demand are obtained by summing the sectoral results (i.e. a 'bottom-up' approach).

Similarly to Table 12, Table 13 presents the 2007-based forecasts together with the most recent data and projections (for EU-25).

Output in agriculture and mining was expected to be weak over the period 2006-08, while output in these sectors over the medium-term (2006-15) was expected to be weaker still, or even to decline. Output growth in agriculture was in fact one percentage point weaker than expected for the period 2006-08, while growth in mining and quarrying output was much higher, partly owing to high commodity prices and strong growth in coal and oil and gas output in 2008. The current expectations for 2006-15 are broadly similar to the 2007 projections, although growth is projected to be weaker in agriculture, while a smaller decline in mining and quarrying is expected.

The 2007 projections for output in manufacturing were too optimistic overall, for both time periods. Output in food, drink and tobacco was particularly weak in 2008, bringing the average annual growth rate for the period 2006-08 down (large falls were experienced in Finland, -5.9%; Estonia, -6.4%; and Slovenia -

11.6%). The broad 'rest of manufacturing' sector (mainly basic industries) also saw particularly weak growth over the period 2006-08, contrary to what was expected in 2007. It is now apparent that these sectors were particularly affected by the recession and reduction in world trade. The more high-skilled engineering sectors were less affected, at least up to 2008. Manufacturing output expectations for the period 2006-15 are now between 0.6 and 1.8 percentage point lower than the previous forecast, mainly due to declines in output in 2009 due to the global recession and weak growth expected in 2010 and 2011.

Table 13. **Sectoral projections of economic output, EU-25**

| | Average annual output growth, 2006-08 | | Average annual output growth 2006-15 | |
|---|---|---|---|---|
| | Expected then | Actual | Expected then | Expected now |
| **Primary sector and utilities** | | | | |
| Agriculture, etc. | 1.6 | 0.6 | 1.4 | 0.9 |
| Mining and quarrying | 0.4 | 3.4 | -1.1 | -0.7 |
| Electricity, gas and water | 0.2 | 0.6 | 0.1 | 1.8 |
| **Manufacturing** | | | | |
| Food, drink and tobacco | 3.6 | 0.4 | 3.7 | 1.9 |
| Engineering | 3.5 | 3.0 | 3.4 | 2.1 |
| Rest of manufacturing | 2.9 | 0.1 | 2.4 | 1.8 |
| Construction | 2.7 | 1.1 | 2.8 | 1.4 |
| **Distribution and transport** | | | | |
| Distribution and retailing | 3.2 | 2.3 | 3.4 | 2.0 |
| Hotels and catering | 0.6 | 2.2 | 1.7 | 1.3 |
| Transport and telecom. | 1.3 | 3.0 | 1.5 | 1.4 |
| **Business and other services** | | | | |
| Banking and insurance | 1.5 | 3.7 | 1.7 | 0.8 |
| Other business services | 3.8 | 3.5 | 4.1 | 2.1 |
| Miscellaneous services | 2.1 | 1.2 | 1.8 | 1.1 |
| **Non-marketed services** | | | | |
| Public administration and defence | 2.2 | 1.9 | 2.4 | 1.4 |
| Education | 1.6 | 1.2 | 1.6 | 1.5 |
| Health and social work | 1.9 | 1.8 | 2.1 | 1.5 |
| All industries | 2.7 | 1.9 | 2.7 | 1.7 |

*Source*: Wilson and Pollit (2011).

For the construction sector, expectations for both time periods were again optimistic compared with actual growth between 2006 and 2008 and the current projections for 2006-15, since it was not known at the time how severe the impact of the financial crisis on the sector would be. The expectations for growth over 2006-15 have been revised down and are now more in line with expectations of growth in the economy overall.

The short-term projections made in 2007 for the hotels and transport sectors were for quite weak growth, reflecting previous years' patterns in some larger Member States. However, the outcome in both sectors was growth of 2-3% p.a. (falls in household expenditure generally came later in the crisis). Output in distribution and retail was roughly one percentage point p.a. less than expected.

Output growth projections for the distribution and transport sectors over 2006-15 have been revised down, reflecting the negative impact of the recession in 2009, and subsequent low growth rates in household expenditure during the recovery. Average growth for the distribution and transport sectors is expected to be broadly similar to the average EU-25 output growth rate for the period.

The banking and insurance sector was hit particularly hard by the global downturn. It comes as no surprise then that medium-term growth expectations have been revised down since the projections made in 2007, from 1.7% p.a. to 0.8% p.a. However, the short-term projections made in 2007 were too pessimistic, as 2006-08 largely represented the final months of booming output before the crisis hit. High growth was observed in the sector in 2007, and growth was particularly strong in eastern and southern EU Member States, such as the Czech Republic, Estonia and Spain; many of which are the countries that were later to be affected most by the crisis.

The other business services sector was also a major casualty of the financial and economic crisis. The short-term projections made in 2007 for average growth were very similar to the actual growth rates over the period 2006-08. However, average annual growth for 2006-15 has been revised down by two percentage points. Countries such as Germany, Spain and France are forecast to see particularly low growth, bringing the EU-25 average down. Furthermore, negative growth in 2009 and low growth during the 2010/11 recovery contribute heavily to the overall low average over the period.

The previous short-term projections for output in miscellaneous services and the public sector were not far off the actual outcomes. However, output growth is now expected to be slower over the period 2006-15, particularly in the public administration and defence sector, in which public spending cuts will have an adverse effect on output growth.

Important factors which may affect rates of employment growth are wage rates and working hours. In the original 2007-based projections it was assumed that, within each sector, nominal wage rates would grow in line with nominal productivity and average working hours would remain unchanged.

It turned out that both these factors help mitigated the effects of the economic and financial crisis, although mainly in the period after 2006-08. For example, many companies chose to cut working hours rather than reduce headcount employment and there were stories of voluntary pay cuts being taken with the same effect. In many ways this was a new phenomenon, which reduced the negative impacts of the recession, but could also be expected to reduce the speed of recovery in employment (as wage rates and working hours would be expected to rebound when growth returns). This is to be considered in the 2011 projections, when more information about trends in recovery should be available.

Finally, the population projections that were used changed between the 2007 and 2010 forecasts; although this primarily affects labour supply, this is expected to be translated into changes in demand through adjustments in long-term wage rates.

Also relevant to the long-term trends in employment is the rate of technological development. In the E3ME model, this is represented by accumulated capital and R&D spending, with the baseline showing an extrapolation of previous trends. The impacts of technology on sectoral employment may be either positive or negative, as machinery may be labour saving, but may also require skilled operators or designers.

The impact of the crisis on technological development remains unclear; credit restrictions may have affected innovation at start-up firms but low rates of return on financial assets may have stimulated private investment.

More time is required for data to become available with which to carry out an assessment of the predicted trends for technological development.

The 2007-based forecasts predicted that employment growth would average 0.7% p.a. over 2006-08 and that productivity would grow by around 2% p.a. However, the crisis meant that this two-year period was anything but standard, incorporating the first part of the recession, when output began to fall, but before job cuts were made. Consequently the annual increase in productivity appears low, at less than 0.5% p.a., and increases in employment were higher than forecast. The links between the output and employment projections are discussed further below.

The differences between the 2007-based forecasts for employment growth over 2006-08 and the outturns are greatest for the primary and utility sectors, in which employment did not decline by as much as expected. Employment in mining and quarrying in fact increased on average each year by a mere 0.4%, partly driven by high commodity prices. The short-term projections for the manufacturing sectors were quite close to the actual outcome, with the biggest difference in the engineering sector, where growth was particularly strong in Estonia, Hungary and Poland. Construction employment also grew by more than was expected.

Table 14. **Sectoral projections of employment growth, EU-25**

| | Average annual growth, 2006-08 | | Average annual growth, 2006-15 | |
|---|---|---|---|---|
| | Expected then | Actual | Expected then | Expected now |
| **Primary sector and utilities** | | | | |
| Agriculture, etc. | -2.8 | -0.6 | -2.1 | -1.8 |
| Mining and quarrying | -3.2 | 0.4 | -2.6 | -2.2 |
| Electricity, gas and water | -1.5 | -0.4 | -1.1 | -1.0 |
| **Manufacturing** | | | | |
| Food, drink and tobacco | 0.1 | 0.1 | -0.2 | -1.2 |
| Engineering | 0.5 | 1.3 | 0.0 | -0.9 |
| Rest of manufacturing | -0.3 | -0.1 | -0.3 | -0.9 |
| Construction | 1.1 | 1.8 | 0.6 | -0.7 |
| **Distribution and transport** | | | | |
| Distribution and retailing | 0.5 | 2.0 | 0.6 | 0.4 |
| Hotels and catering | 1.8 | 2.1 | 1.8 | 0.9 |
| Transport and telecom. | -0.1 | 1.3 | 0.0 | 0.2 |
| **Business and other services** | | | | |
| Banking and insurance | -0.1 | 0.8 | 0.0 | -0.5 |
| Other business services | 2.7 | 4.2 | 2.9 | 1.6 |
| Miscellaneous services | 1.5 | 1.3 | 1.6 | 0.7 |
| **Non-marketed services** | | | | |
| Public administration and defence | 0.2 | 0.1 | 0.2 | -0.1 |
| Education | 0.8 | 0.9 | 0.9 | 0.4 |
| Health and social work | 0.8 | 1.7 | 0.8 | 0.9 |
| All industries | 0.7 | 1.5 | 0.7 | 0.2 |

*Source*: Wilson and Pollit (2011).

For the distribution and transport sectors the actual average annual growth rate over the period 2006-08 was higher than projected in 2007. For all of these sectors high employment increases were mainly driven by higher than expected employment growth in central and eastern European countries. For example, employment in distribution and retailing increased by 8.0% p.a. in Lithuania, whereas the 2007 projections had only predicted a 3.8% p.a. increase. Other EU-15 Member States also saw large increases in employment in some of these sectors, such as the hotels and catering sector in Ireland, and the transport and telecommunications sector in Spain.

The short-term annual employment growth projections made for business services were also lower than the actual outcome, by around 1-1.5 percentage points. Nevertheless, the projections had correctly predicted that this sector would see the strongest employment growth over the period. The short-term predictions made for employment in miscellaneous services and the public sectors were broadly similar to the actual outcomes, although lower for health and social work.

As for the medium-term projections, the current expectations for the period 2006-15 have not changed a great deal for the primary and utility sectors, transport and telecommunications and the health and social work sector, compared to the projections made in 2007 (although the most recent projections anticipated public spending cuts, in most cases they were too early to incorporate specific packages).

However, annual employment growth in all other sectors is now projected to be lower than previously thought, by between 0.2 and 1.3 percentage points. Overall, this translates into average annual employment growth of 0.2% over the period, compared to the more positive 0.7% growth in the 2007-based forecast. The differences reflect the downturn in 2009-10, expected impacts on government balances and the revised demographic projections.

#### 8.4.3. Differences between output and employment projections

The aim of this evaluation exercise is to identify where discrepancies between predicted and actual outcomes occurred. So far the focus has been on projections of output, which were an input to the forecast, and the results for employment. However, of most interest is the process of going from output to employment, which is the focus of this section. Table 15 presents a clearer picture of the differences in the projections made in 2007 and the actual short-term outcomes and current medium-term projections. Differences are shown for six broad sectors.

The most obvious point to note from Table 15 is that output growth proved to be weaker than forecast whereas the growth in employment proved to be stronger than forecast. This outcome can be explained by the short two-year period used and the position it had in the economic cycle; by the end of 2008 economies were contracting but, as changes in employment tend to lag changes in output, at this point there were no notable impacts on employment. Hence the implied error in productivity (1.6% annual growth) is due to short-term reductions in aggregate demand.

It is difficult to infer much for the sectoral level. The sectors with the largest underestimates of employment were also those that grew over 2006-08 and the sectors that were hit hardest by the crisis (manufacturing and construction) are the ones with the largest differences. There is some correlation between the output and employment errors.

It is easier to present similar patterns between the 2007 and 2010 medium-term projections. In general, half of the reduction in the forecast for output growth rates translates into a reduction in employment, which means in effect that the loss of output is split 50/50 between loss of productivity (i.e. wages) and loss of employment levels by 2015. The sectors with significant variation from this central trend are mainly those which have seen the largest changes in employment, notably construction and some manufacturing sectors. The public sectors are the least elastic in this respect, with lower levels of output translating less into lower levels of employment.

Table 15. **Differences between output and employment projections (percentage points)**

| | Difference in 2006-08 actual outcome to projections | | Difference in sets of 2006-15 projections | |
|---|---|---|---|---|
| | Output | Employment | Output | Employment |
| Primary sector and utilities | 0.2 | 2.3 | 0.7 | 0.3 |
| Manufacturing | -1.8 | 0.4 | -1.1 | -0.8 |
| Construction | -1.6 | 0.7 | -1.4 | -1.3 |
| Distribution and transport | 0.3 | 1.2 | -0.8 | -0.2 |
| Business and other services | 0.4 | 0.9 | -1.6 | -1.0 |
| Non-marketed services | -0.2 | 0.4 | -0.6 | -0.2 |
| All industries | -0.8 | 0.8 | -1.0 | -0.5 |

*Source*: Wilson and Pollit (2011).

### 8.4.4. Supply of labour

In this section the projections for growth in the EU-27 labour force and participation rates that were made in 2008 are compared with the most recent ones made in 2010. A comparison of population projections is not shown, since these have not changed (they did, however, change between the 2007 demand and 2008 supply forecasts).

As there is only one year of difference between the two forecasts, it is not possible to compare the predictions with actual outcomes, so this section focuses upon the medium-term forecast.

Compared with employment demand, the effects of the financial and economic crisis are less important to the forecasts of supply (i.e. demand falls by more than supply so unemployment increases), but there are impacts, for example:

(a) lower output growth and wage rates mean people may be less inclined to seek work;
(b) higher unemployment causes some people to withdraw from the labour market.

There is another important factor, changes in migration patterns, that has not been considered as it is implicitly included in the exogenous population projections and so is outside the scope of our assessment.

Table 16 shows the differences between the predictions for participation rates in 2020 and for the growth in the labour force by 2020 compared with 2008. A large part of the differences in results is due to the extra year of data being available for the later forecast; the differences in some age groups were quite large, possibly indicating a response to the crisis but also possibly variation within the LFS data.

The similarity between the two forecasts for 2020 is also a reflection of the methodology that was used to produce them. The 2009 forecast and 2010 update essentially uses the 2008 projections as a starting point, plus an additional year of data and a representation of the recession. However, as recovery is well under way by 2020, the main differences are due to the additional year of data. While this means that there is a degree of consistency between the two forecasts, it cannot be said that one is a confirmation of the other, and further assessment should be made.

This section has attempted to identify and isolate differences between predicted and actual (and revised predicted) outcomes for employment. The differences have been split into differences in inputs (mainly economic projections and demographics) and differences in processes within the E3ME model.

In summary, it is difficult to draw many firm conclusions from the evaluation so far. This is partly because only a short period has passed since the original projections were made, and partly because this period was dominated by the economic and financial crisis, which was not included in the baseline economic (and therefore employment) projections used in 2007. The level of uncertainty is still rather high as different paths to recovery are considered and concerns remain over levels of consumer and government debt in certain parts of Europe.

Table 16. **Labour force and participation rate projections, EU-27**

| | Participation rate projections | | Labour force projections | |
|---|---|---|---|---|
| | 2020 participation rates expected in 2008 | 2020 participation rates expected in 2010 | 2008-20 % growth expected then | 2008-20 % growth expected now |
| Male 15-19 | 0.23 | 0.22 | -18.4 | -22.0 |
| Male 20-24 | 0.69 | 0.65 | -12.1 | -14.3 |
| Male 25-29 | 0.89 | 0.87 | -6.0 | -6.6 |
| Male 30-39 | 0.95 | 0.93 | -2.4 | -2.9 |
| Male 40-49 | 0.94 | 0.93 | -2.2 | -2.5 |
| Male 50-54 | 0.87 | 0.88 | 10.7 | 12.2 |
| Male 55-59 | 0.72 | 0.76 | 14.8 | 19.7 |
| Male 60-64 | 0.42 | 0.45 | 31.2 | 31.1 |
| Male 65+ | 0.08 | 0.08 | 47.9 | 34.8 |
| Female 15-19 | 0.21 | 0.20 | -13.4 | -19.7 |
| Female 20-24 | 0.57 | 0.56 | -13.3 | -14.3 |
| Female 25-29 | 0.80 | 0.79 | -3.7 | -3.4 |
| Female 30-39 | 0.79 | 0.80 | -2.9 | -1.3 |
| Female 40-49 | 0.81 | 0.82 | -1.2 | -1.1 |
| Female 50-54 | 0.72 | 0.74 | 10.6 | 12.1 |
| Female 55-59 | 0.56 | 0.59 | 21.9 | 24.8 |
| Female 60-64 | 0.24 | 0.25 | 32.8 | 29.2 |
| Female 65+ | 0.03 | 0.03 | 36.3 | 18.9 |
| Total | 0.56 | 0.56 | 1.1 | 1.3 |

*Source*: Wilson and Pollitt (2011).

At the moment, several conclusions can be drawn, though, which are outlined below. Although the conclusions at this stage are rather vague and tend to focus on inputs used rather than the means with which the projections are used, there are some important messages to be nonetheless drawn:

(a) the medium-term projections must take into account short-term fluctuations, as showed in the period 2006-08. In particular, dynamic factors, with employment often lagging output, should be considered;

(b) revisions to data can have a substantial effect on the projections, even if other factors remain unchanged. It is important to use the most recent data;

(c) government policy, which is treated as exogenous in the modelling, plays an important role, particularly in periods where ‘business-as-usual’ cannot be assumed. Most obviously affected are high-employment sectors such as health and education, but there are other less obvious examples, such as the motor vehicles sector that benefited from car scrappage schemes in Europe.

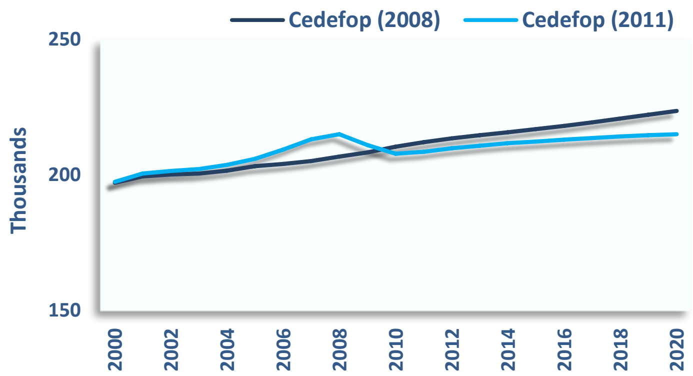
#### 8.4.5. Evaluating the occupational employment projections

This section moves on to consider the implications for the demand for skills, focusing on occupational employment projections, based on the sectoral projections produced using E3ME.

The emphasis is on the results over a 5-10-year forecasting horizon. The assessment begins with a qualitative review of the projections, comparing the predictions with the outcomes, as reported in subsequent reports.

Figure 8 illustrates several issues raised in the earlier discussion quite neatly. It compares the latest 'quick update' (used in Cedefop, 2011) results with those from the pilot demand projections conducted in 2007 and publish in Cedefop (2008).

Figure 8. **Comparison of aggregate employment trends**

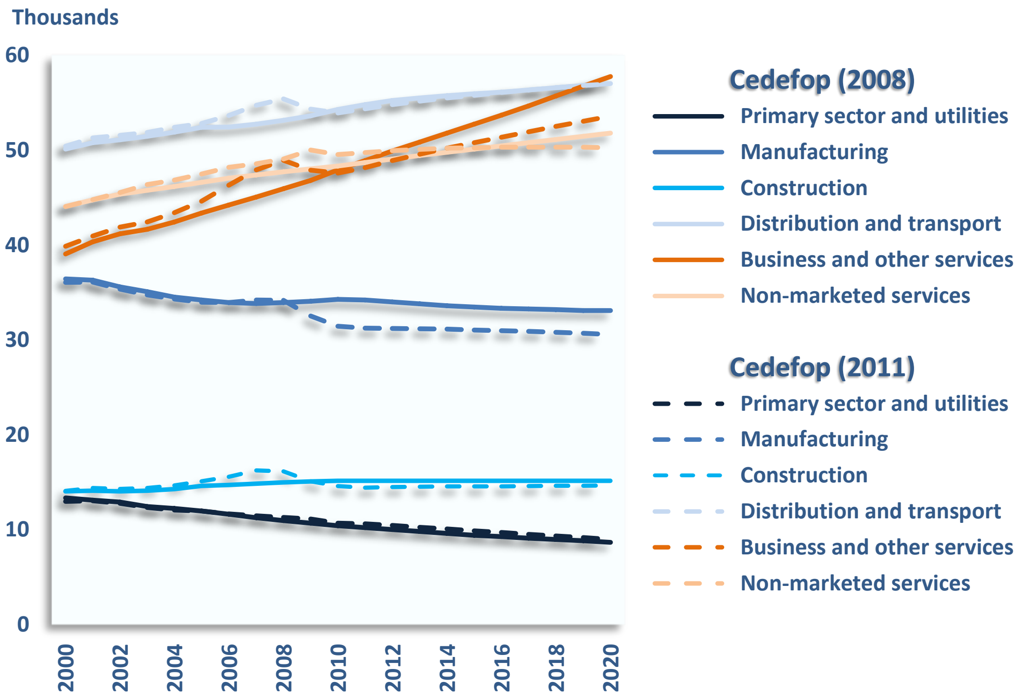



*Source*: Wilson and Pollitt (2011).

The results show that historical data revisions dominate the differences up to 2008, while the financial crisis and subsequent worldwide recession explain the main differences in the post-2008 period. Although few financial commentators predicted the financial crash, this kind of 'singularity' is almost impossible to build into conventional economic/labour market assessments, and the events of 2008 were not predicted by any mainstream forecasters.

The important question is how much does this matter in terms of the general messages emerging from the Cedefop projections. Figures 9 and 10 suggest that it is not as important as might first be thought. Although it is clear that Cedefop projections based around the E3ME macroeconomic model were not able to predict the crisis, many of the patterns in terms of industrial and occupational employment that it provided remain surprisingly robust.

Figure 9. **Comparison of aggregate employment trends by sector (thousands of jobs EU-27+)**



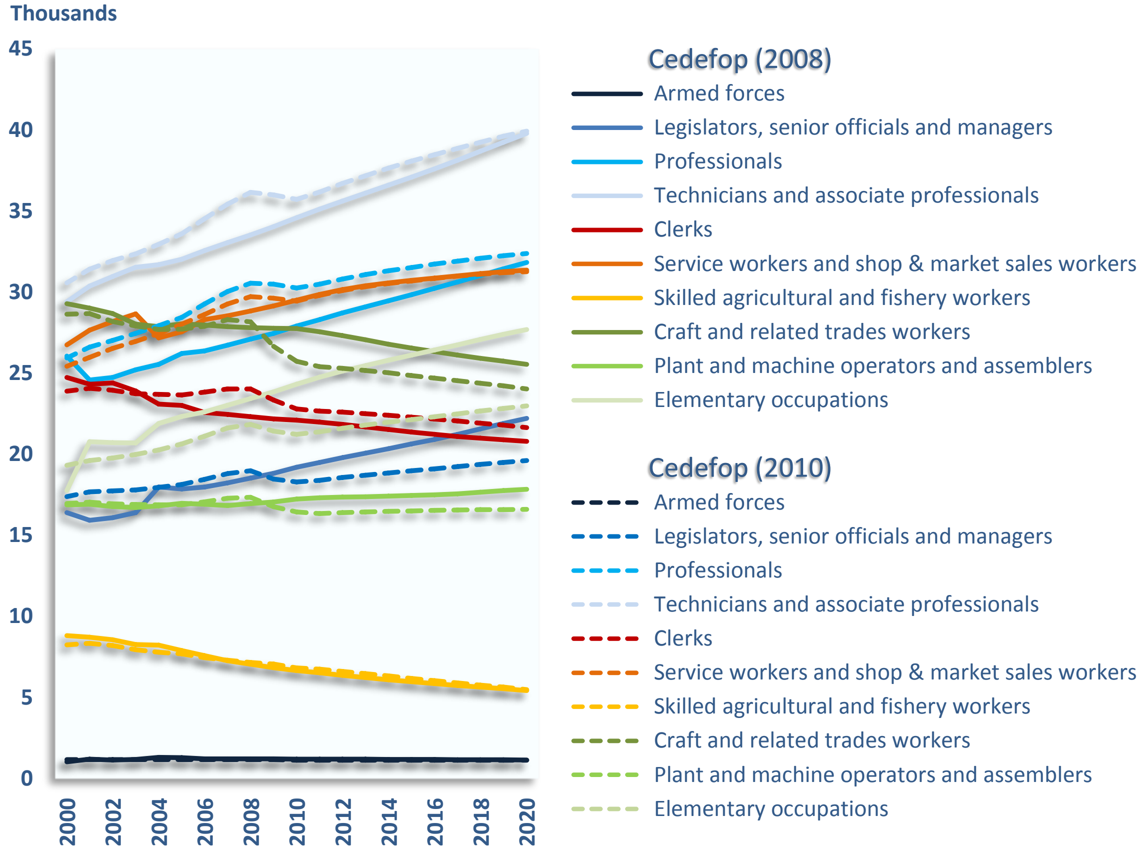

*Source*: Wilson and Pollitt (2011).

Figure 9 illustrates the time series profiles for employment by sector in the two sets of projections (Cedefop, 2008; 2011). Although the impact of the crisis and recession is immediately apparent again in most of the sectoral profiles, it is also clear that the relative sizes and underlying trends in the two sets of projections are very similar.

This is probably even more the case if a comparison is made of the occupational or qualification trends as shown in Figures 10 and 11. While there are obvious differences between the two sets of projections, they are again dominated by historical data revisions and the impact of the crisis.

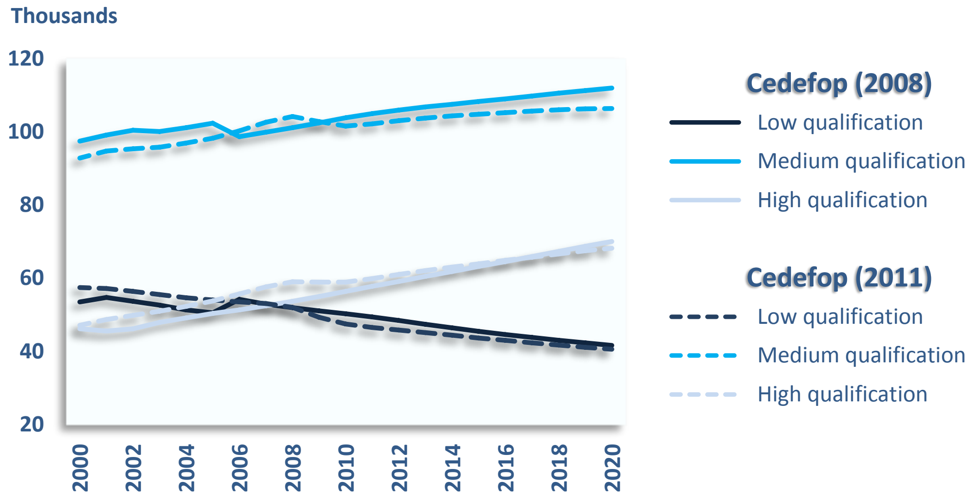
The broad patterns in the historical data and in the projections remain very similar in the two sets of projections. This is even more true if the focus is on replacement needs and total requirements as opposed to net change (expansion demands), as the latter are highly dependent on the level of employment as opposed to the change over time.

Figure 10. **Comparison of aggregate employment trends by occupation**




*Source*: Wilson and Pollitt (2011).

Figure 11. **Comparison of aggregate employment trends by qualification**




*Source*: Wilson and Pollitt (2011).

CHAPTER 9.
# Qualitative evaluation

The quantitative evaluation of the forecasts' accuracy was accompanied by the qualitative survey of the usage of the skill forecasts and related analyses. The main purpose of such an evaluation was to obtain information on the meaningfulness of the forecasting efforts and to steer the project towards meeting the needs of different stakeholders better. The aim was not only to gather information on general satisfaction with the presentation of the results and documents, but also to assess their quality and further use. Furthermore, we have examined the perceived quality of the forecast and explored future directions.

This chapter is based on the background paper by Kriechel and Wilson (2011a). Section 9.1 introduces the survey methodology, including the main idea of the evaluation process. Section 9.2 presents results on the general content and use of the skills forecast. Section 9.3 assesses the quality of the forecast. Section 9.4 gathers ideas for further development of the survey.

## 9.1. Methodology

An Internet-based survey containing 16 questions (Annex 4) was used. Respondents could answer these questions in less than 10 minutes. The respondents received an invitation e-mail and an endorsement e-mail from Cedefop. All respondents that did not answer within 10 days received a reminder e-mail.

Ideally, the population of respondents would consist of all users of the skills forecast. However, since there is no close monitoring of the user base, we approximated the population by inviting all country experts participating in workshops on skill forecasting organised by Cedefop and all other experts who requested access to the detailed country results. In addition, participants in the Agora 2009 and Agora 2010 conferences were also invited to take part. We expected that the members of these groups should have, at least through their participation, some knowledge of skill forecasting methodology and outcomes as they were key topics of the workshops and conferences. The survey was conducted by ROA based on the contact list provided by Cedefop. After eliminating the duplicate addresses of persons that participated in several events, we were left with a base population of 255 people. They all received an invitation letter from Cedefop, followed by an invitation e-mail that provided a link to the Internet questionnaire. There were 102 respondents, yielding a raw response rate of 40% (102/255).

Table 17 summarises the information sources of the respondents. More than 90% of the respondents have received at least the synthesis report on the project, while about three-quarters of the respondents replied that they had received the background technical report and the briefing notes.

Table 17. **Information sources of respondents**

| | Not received /don't know this | Received but haven't read it | Received but found it not interesting | Received and found it somewhat interesting | Received and found it very interesting |
|---|---|---|---|---|---|
| Briefing Notes | 23.08 | 6.15 | 3.08 | 29.23 | 38.46 |
| (Synthesis) Reports | 10.77 | | 7.69 | 23.08 | 58.46 |
| Background technical report | 27.69 | 9.23 | 1.54 | 32.31 | 29.23 |
| Observations | 65 | | | | |

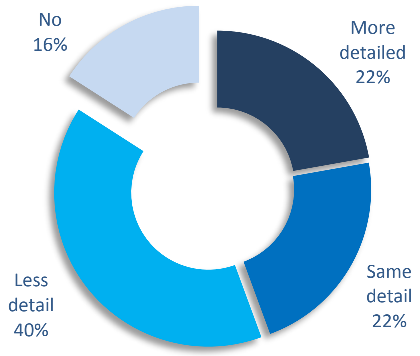
*Source*: Online survey, ROA, 2010.

As expected, the most important information source was the synthesis report. All respondents who reported having received it had also read it and found it interesting. The background technical report was more often not read. This could be due to the technical nature of the report, which appeals only to readers that try to understand the underlying mechanisms of the modelling framework. An interesting question is whether readers who were not entirely content with the background technical report, i.e. those who found it ‘somewhat interesting’, would have liked to have more technical detail or whether they found the reports too complicated. However, we did not investigate this issue in the survey, so we cannot make a distinction here.

As Figure 12 shows, the European skills forecasting provides policy-makers and experts in the fields either with an additional skills forecast on a comparable (22%) or inferior level of detail (22%), but for other countries the EU skills forecast is the only one (16% of the cases) or a more detailed skills forecasting model (40%) than what is nationally available.

Most respondents of our population were, as expected, involved in some form of skills forecasting. Inevitably, some respondents of the survey were country experts who are part of the consortium developing and making the forecast.

Figure 12. **Availability of national skills forecast**




*Source*: Online survey, ROA, 2010.

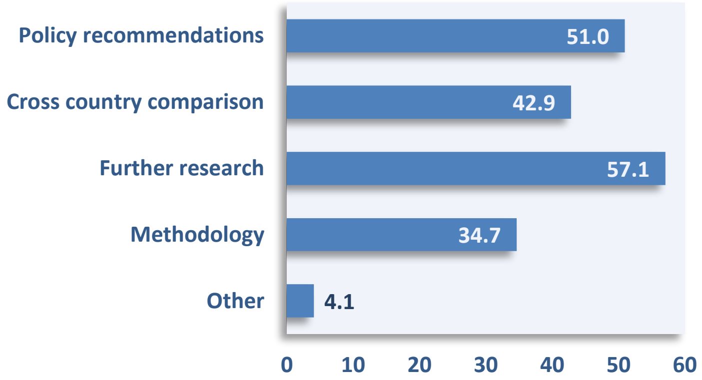
## 9.2. Contents and use of skills forecasting results

The project's aim is to develop and publish regular medium-term skill forecasts at pan-European level. In addition to developing a European skill forecast, it should also provide individual country forecasts using a consistent and comparable methodology so that country results can be compared across Europe. The expertise built up and shared with country experts could provide a foundation for the development or improvements of national skill forecasts.

In the question on the main use of the Cedefop skill forecast to the respondent, our aim was to determine how the project helps users. The first two aspects in Figure 13, namely policy recommendation (51%) and cross-country comparison (43%), are part of the main goal of the project, which is to provide skills forecast to support policy recommendations, on both the EU and national comparative levels. Of the respondents, 57% also replied that the project is an important source for use in further research, while 35% reported using methodological aspects of the project.

The results of the skills forecast are presented in Excel workbooks. The workbooks provide detailed results on all aspects of the model, next to the underlying data fed into the forecasting model. It provides a convenient way to generate tables and figures to the user's specification. In addition, the workbook allows some changes in the underlying data and assumptions of the model. The effects of these changes applied by the users are then calculated within the workbook. This allows the user to make modifications and carry out interactive analysis.

Figure 13. **Main use of forecasting projects result**




*Source*: Online survey, ROA, 2010.

Table 18 summarises the respondents' evaluation of the two elements of the workbook: the ability to change underlying data and the availability of detailed (underlying) data for the user. Especially the availability of the detailed data seems to be important to the respondents. Of the respondents, 52% reported this to be very important and an additional 39% as important. Being able to change the underlying data or assumptions of the model is also appreciated: almost 20% find this very important and an additional 48% find it important. Thus, the availability of the detailed data seems to be the more important aspect that almost all users want, while a smaller group actively use the possibility to change underlying data. The detailed workbooks were evaluated as important, and given the policy recommendations made and the additional research undertaken at national level using the outcomes and data of the skill forecasting project, it would appear sensible to continue providing detailed data to national experts.

Table 18. **Workbook usability**

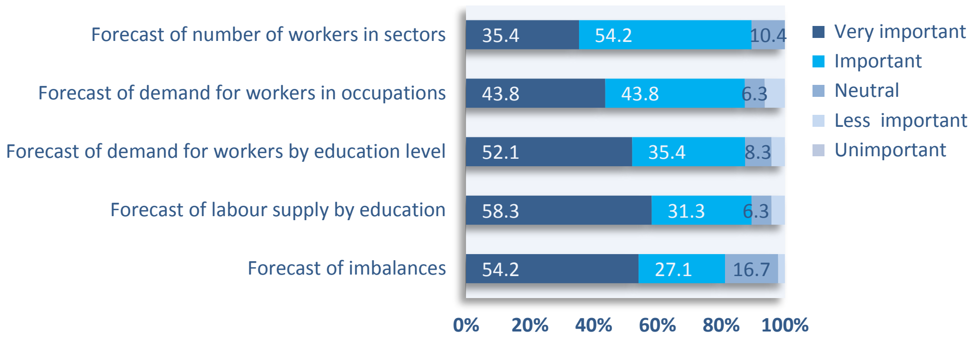
| | Able to change underlying data | Detailed data in workbooks |
|---|---|---|
| (1) very important | 19.57 | 52.17 |
| (2) important | 47.83 | 39.13 |
| (3) neutral | 21.74 | 4.35 |
| (4) less important | 2.17 | 4.35 |
| (5) not important | 8.70 | 0.00 |
| Observations | 46 | |

*Source*: Online survey, ROA, 2010.

The importance of the forecasting elements is reported in Figure 3. A distinction is made between the forecast of the number of workers in sectors, the demand for workers by occupation, the demand for workers by education level, the supply by education level, and of the imbalances resulting from supply and demand forecasts.

Figure 14 shows that all aspects are considered important or very important by most respondents. No aspect has less than 80% of respondents reporting less than important. Ordering the results by the 'very important' elements yields the following order: most important is the forecast on supply by education (58%), the imbalance forecast (54%), followed by the forecast of demand by education level (52%) and by occupation (44%), with the forecast of the number of workers in a sector (35%) in last place. However, this picture changes if the 'very important' are combine with the 'important' responses. Then the forecast of the number of workers by sector comes out on top, together with the forecast on labour supply by education (both 90%).

Figure 14. **Importance of forecasting elements**



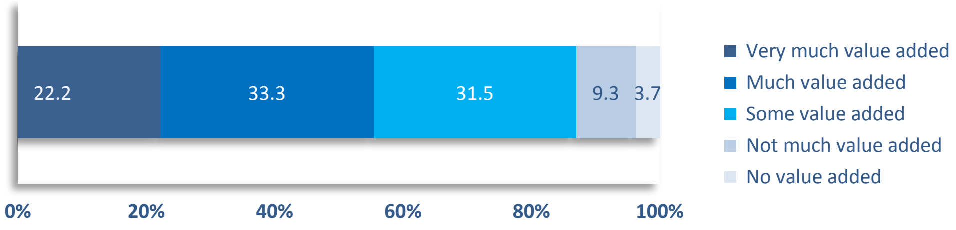

*Source*: Online survey, ROA, 2010.

## 9.3. Qualitative evaluation of results

How successful has the project been so far in generating usable and plausible results at the national and cross-country levels? The overall evaluation of the forecast, as given in Figure 4, shows that the respondents see sufficient value added by the project. Twenty-two per cent think that very much value is added, followed by 33% who see much value added and another 32% who see some value added. Only 13% of the respondents see not much or no value added by the forecast.

We also asked respondents to evaluate the results of the individual elements relative to national results if they existed, or their expectations of results if no national results existed. The plausibility of the results is evaluated favourably for all aspects of the model: demand by sector, demand by occupation, replacement demand, demand by qualification and supply of skills. While there are slight variations in the mean outcome of these five aspects, none are statistically significant. The examination of confidence band gives some reassurance that the respondents evaluate the forecasting outcomes as plausible. This is confirmed in Table 19, which summarises the results from a question on the overall evaluation of the forecast plausibility, including all elements. More than half of the respondents report that the results are either close to national forecasts, or go in the same direction.

Figure 15. **Overall evaluation of forecast (% of total)**




*Source*: Online survey, ROA, 2010.

However, about one third of the respondents report that the outcomes partially contradict national outcomes. This is to be expected. The data and methodology used are geared towards a unified framework for all countries. Especially in countries that have national forecasting models, it is likely that these are more optimised to the national data and include much more knowledge of national institutions and other elements influencing projection outcomes. In this project, we attempt to find those contradictory outcomes by discussing results with country experts and asking for their opinions on the plausibility of the results already during the forecasting process.

Table 19. **Overall evaluation of forecast plausibility**

| | |
|---|---|
| Don't know | 10.9% |
| Very close to national forecast/ expectations | 6.5% |
| Same direction of outcomes | 50.0% |
| Partially contradicting outcomes | 30.4% |
| Contradicts national forecasts/ expectations | 2.2% |
| Observations | 46 |

*Source*: Online survey, ROA, 2010.

The overall evaluation of the importance of the project and of the plausibility is thus quite favourable. However, we should not rest on our laurels, but should rather look into improving the current model and, whenever possible, including and resolving the contradictions in the outcomes of national and the EU forecasts.

## 9.4. Further development

The project has developed a well-functioning model of skills forecasting. While there are many aspects that can be improved or extended, given the results described in the previous section, we have already achieved an important building block of the project. The model provides forecasts that country experts evaluate as plausible and largely in line with national forecasts or expectations. Further development work should go into scenario development, both on a country or sectoral scale. These could be included in the presentation of the findings, for example by allowing the user to change some key assumptions in the workbooks. The wish for more detailed skills forecasting needs to be supported by the data availability.

Table 20 summarises the outcome of the question on suggested directions for further development. Scenarios both on the country level (50.9%) and on the sector level (54.5%) score high on the respondents' wish list. In addition, more detailed skills and competence forecasts are wished for (54.5%). Expanding the forecast by including fields of study might already be a step into that direction. However, more detailed forecasts hinge upon the availability of reliable data.

Table 20. **Suggested directions of further development**

| Direction of further development | |
|---|---|
| Policy simulation | 41.8% |
| Technological/ organisational changes simulations | 29.1% |
| Individual country scenarios | 50.9% |
| Individual sector scenarios | 54.5% |
| More detailed skills/ competences forecast | 50.9% |
| Other | 3.6% |
| Observations | 55 |

*Source*: Online survey, ROA, 2010.

The survey revealed that the EU-wide skill forecasting project provides many countries with skill forecasting information which was not previously available. The forecasts are used for policy recommendations, to compare developments in different countries, and to provide the basis for further research on skill demand and supply. National experts see the results of the skill forecast as more plausible than national forecasts or national expectations of future developments. Almost 60% of the respondents report that the overall results are close to national forecasts or expectations, whereas one third report (partial) contradictions between the forecast and national sources.

# CHAPTER 10.
# Summary and conclusions

## 10.1. Summary

The labour markets of the EU and the individual Member States are undergoing structural changes. The shifts towards a more service-oriented and knowledge-based economy are changing the structure of labour demand, mainly in terms of qualifications required. Ageing means that the labour force must be used more efficiently. Austerity measures will reduce public investment in general, including education. Europe cannot afford to waste the potential of the labour force through unemployment and it needs to ensure efficient use of public investment in education. Anticipation of skills supply and demand can provide crucial information about long-term trends and thus support decision-making.

Cedefop's skills supply and demand forecast is the only European-level exercise based on comparable data. It relies on up-to-date and sophisticated methods and is constantly being developed with the help of key European research institutes.

Anticipation of future skills needs and supply is also taking place in Member States. Cedefop does not intend to compete with or replace efforts at the national level. With this publication, Cedefop intends to present a general overview of the methodology and to inspire those engaged in forecasting exercises or trying to find relevant background material to start their own forecasts.

A consistent and comparable data base is a prerequisite for obtaining reliable models and results. For this purpose, several data sources needed to be adopted: it is necessary to link labour force survey (LFS) data, which provides desegregations by occupation and qualification, and national accounts (NA) data, which provide complete estimates of sectoral employment consistent with sectoral value added and other macroeconomic indicators.

The Cedefop methodology (Figure 1) is based on econometric modelling with certain elements of input-output techniques. The general framework is built on a modular approach, which allows for independent development, fine-tuning and the extension of particular parts in time. The core element is a multisectoral macroeconomic model which produces labour-demand data by sector (E3ME). Its augmented version also produces labour market participation rates for labour supply data. The labour demand by sector is further processed in modules of labour (expansion) demand by occupation and qualification. This forecast also uses a replacement demand module to produce a forecast of those leaving the occupation for different reasons (retirement, mobility, etc.). By combining expansion and replacement demand, the future total job openings can be calculated. Stock and flow modules produce results for labour supply by qualification. A combined stock-flow model would be ideal but is not currently feasible given the lack of data. As labour supply and demand are interrelated, the algorithms to calculate imbalances are not straightforward. Cedefop developed a module on imbalances to deal with this issue.

In Cedefop's forecast the E3ME provides the links between the labour market and the wider economy. The module is made up of four sets of equations: employment demand, average wages, average hours worked and participation rates. E3ME, which provides a coherent European perspective, allows for the production of different projection scenarios for labour supply and demand.

The supply of skills module is based on demographic forecasts and assumptions for individuals to acquire certain level of qualification. Ageing, which is generally spread over European labour markets, will result in decline of younger cohorts in the population and in the labour force. A decline in the number of young people in the labour force is also expected because younger people are staying longer in education than in the past. At the same time, older people usually stay in employment longer due to legislative changes in the age of retirement. Based on such assumptions, the stock of qualifications can be relatively easily derived. It is more complicated to follow the flows among different age categories and age groups important for a more plastic view of the structure of labour supply. A stock–flow model would be ideal for forecasting supply of skills. Producing such a model for the qualification mix of individuals across Member States appears feasible, but is not without problems. The EU-LFS data are now available over a sufficient time period to construct pseudo cohort information that allows modelling of the transition from school to work. But when modelling qualification levels, stock-flow modelling has to deal with the issue of transitions between education pathways. It has been demonstrated that while primarily concentrated among younger individuals, they still cut across a significant age range of at least 16 to 30 years. There are several other minor issues in the modelling process that can be dealt with in a fairly straightforward manner. These include the absence of information on qualifications for those above retirement age and not in employment, and the possible lower bounds for proportions of the population with low qualifications. Other more demanding issues relate to modelling explicitly the effects of emigration and immigration on the qualification mix in different countries. Nevertheless, the present study demonstrates that appropriate data are not available for such modelling at pan-European level.

Changes in the level of employment (expansion demand) disaggregated by occupation and qualification are calculated in two modules (EDMOD and QUALMOD). The projections of employment are derived by mapping projected shares of occupations and qualifications to sectoral employment projections obtained from E3ME. These results incorporate the latest data and, as far as possible, consider latest global developments. In recent years, Cedefop has explored greater level of detail in various data sets, including trials to extend the set of results by more detailed levels of educational attainment and attempts to incorporate fields of study. Given the exploratory nature of the most recent work, not all of the issues and problems have been resolved yet. Although present results are encouraging, so far the approach has not offered a viable alternative to simpler methods of developing projections of changing skill demand patterns based on extrapolative techniques. The conceptual underpinnings of a 'behavioural' theory of labour demand, including occupations and education levels, already exist. This theory has its roots in production and cost functions, which have been adapted to estimate employer demand for labour. Subsequent work has been carried out to extend Cedefop's framework to look at skill-biased technological change (SBTC). The possibility to simultaneously combine estimates of all 81 education/occupation categories was examined. An initial exploration of the data suggests that interesting patterns emerge, but it is still too early to draw conclusions about the success of further work in this area. However, such work would undoubtedly lead to a significant improvement in understanding of recent and current labour demands and employment.

Replacement demand is an important component of the overall labour demand. This component of labour demand represents demand due to people leaving occupations for different reasons who need to be replaced. The cohort component method is used to produce net flows from the labour market. However, the tests of alternative approaches to net flows by occupation estimates can be used only at national level due to the current lack of necessary data sets at European level. We could observe, taking one example for each country, that the flow estimates from panel versus cohort-component approach are quite similar, which leads to conclusion that the cohort component method currently used in Cedefop's forecast is a good substitute for a methodologically superior panel-based approach if the underlying causes of replacement needs do not need to be identified.

Comparing labour demand and supply to derive potential imbalances is not straightforward. In Cedefop's skills forecast, supply and demand are calculated separately using comparable data and assumptions. In presenting the implications for imbalances it is essential to emphasise that both the trends in supply (towards a more highly educated workforce) and the trends in demand (towards greater use of such people in employment) are hard to predict precisely. They are also interrelated (supply can to some extent help to generate its own demand, and demand can also generate supply to some degree) but the models are not able to capture these interactions. Several indicators are currently being developed to improve the link between occupational demand and educational supply. Both the indicator based on education shares in occupations and the RAS-based indicator of constraint and measure of change point in the same direction: lower-level occupations will face more adjustment needs and potentially more difficulties in hiring workers. The indicators of imbalances would benefit greatly if a distinction by field of study could be made in education supply.

Finally, when improving forecasting methods, it is important to look back and assess how accurate and useful past projections have been. The detailed analysis of errors in the projections suggests that the aggregate projections of changes in GDP and total employment in the initial pilot studies (2007-08) were overtaken by events as a result of the financial crisis. Cedefop’s occupational projections appear at least as accurate as the industrial ones, which is somewhat surprising given the quality of the data. However, this is probably (in part at least) due to the relative stability of the occupational structure. As for the industry results, the projections generally mirror the observed directions and magnitudes of change. The levels are less accurately projected but this is linked to revisions of the base year figures and to the impact of the crisis. The direction of change appears to have been correctly forecast by Cedefop.

To examine the accuracy and numerical quality of the forecast is an important issue. The research on usability and overall satisfaction of the results is no less important. A user survey revealed that Cedefop’s skill forecasts provide many countries with information which was not previously available. National experts use Cedefop forecasts mainly for policy recommendations, comparisons of developments in different countries, and further research into skill demand and supply. Moreover, the results are considered plausible if compared to national forecasts or national expectations on future developments.

## 10.2. Continuing dialogue and country expert input

Cedefop’s conceptual framework, related improvements and results have been developed by a team of experts from leading institutions in economic modelling, labour market analysis and skills anticipation. The uniqueness of Cedefop’s approach implies that it is not possible to compare results or methods with similar exercises.

Cedefop has put together a group of experts within its Skillsnet network to ensure the high standard and credibility of the methods used and to validate the plausibility of the results. These experts represent the whole spectrum of professionals starting from practitioners, VET providers through economists and labour market analysts to statisticians and econometricians. Regular workshops, ad hoc discussion groups and informal communication also provide input for further developments and improvements. Without the expert's input and quality checks it would have been difficult to achieve the same outcomes and establish the integrity of the project.

Cedefop is really grateful for each input provided by experts. All other experts and stakeholders are invited to join and provide their own views and suggestions.

## 10.3. The way forward

Important developments have been made since 2005, when the first discussions about forecasting skills supply and demand for the whole of Europe started. Several high-level techniques were put together to provide a unique conceptual framework able to forecast skills needs for different sectors and occupations in all EU Member States. A Thorough analysis of the various difficulties and issues suggests ways to improve and develop. Several avenues have already been explored with more or less satisfactory results. They were reviewed in this publication; many more are, however, still awaiting detailed exploration.

One of the important issues to be explored further is the treatment of both sides of the labour market in terms of qualifications. It is clear that there is a complex relation between the supply of and the demand for qualifications which cannot be easily captured. Sets of indicators to indicate potential labour market imbalances were developed, but more has to be done to test the robustness and reliability of the data they use and their conceptual clarity. This is the focus of upcoming work, which will provide a better picture of the different forms of skills mismatches and imbalances.

Qualitative inputs are also important in forming different policy scenarios. Extending the current general framework by some of the computable general equilibrium techniques will be done to improve the understanding of trends in qualification needs under different (policy) scenarios (e.g. greening of economies, ageing, etc.).

Other methodological developments are awaiting better data. There is a common agreement that internal mobility and migration from and to non-member countries is affecting both the supply and demand for skills. But available information and data about migration do not allow the incorporation of this phenomenon into the overall modelling framework. More disaggregated data on the skill profiles of occupations and sectors are being developed. But they are based on the O*Net, which reflects the situation in the United States and not Europe. The use of these data for building a forecast for Europe is thus questionable.

Finally, the skills forecast is not the only skills analysis activity of Cedefop. Other projects focus on different aspects of skills formation and use in relation to education and the labour market. One of them investigates and analyses skill mismatch in Europe, including new empirical analysis of skills obsolescence based on a Cedefop pilot survey in four countries. Another provides evidence on sectoral skills and qualification requirements and supports anticipation of skill need at European and international level by providing information on new and emerging skill needs in selected sectors. Cedefop also contributes to the current debate on the greening of the economy and related skill needs. The results are used for different analyses such as the polarisation and ageing of occupations. Finally, work is in progress on an employers' survey on changing skill needs. All these activities will form different building blocks to provide better understanding of skill needs and trends in Europe and will also form part of the flagship initiative of Europe 2020 'the agenda for new skills and jobs'.

# References

Acemoglu, D. (2002). Technical change, inequality and the labour market. *Journal of economic literature*, Vol. 40, Issue 1, pp. 7–72.

Barker, T. et al. (2006). Decarbonising the global economy with induced technological change: scenario to 2100 using E3MG. *Energy journal*, Vol. 27, pp. 143–160.

Berndt, E.; Christensen, L. (1973). The translog function and the substitution of equipment, structures, and labour in US manufacturing 1929-68. *Journal of econometrics*, Vol. 1, Issue 1, pp. 81–114.

BLS (1997). *Handbook of methods; Chapter 13: employment projections.* Washington, DC: BLS, Bureau of Labor Statistics.

Bosworth, D. et al. (2011). *Modelling and forecasting the demand for occupations and qualifications*. Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Bosworth, D.; Wilson, R. (2011). *Forecasting skill supply and demand in Europe to 2020: stock flow model and other modelling developments.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Cedefop (2008). *Future skill needs in Europe: medium-term forecast; synthesis report*. Luxembourg: Publications Office. http://www.cedefop.europa.eu/EN/publications/12968.aspx [accessed 7.5.2012].

Cedefop (2009). *Future skill supply in Europe: medium-term forecast up to 2020; synthesis report*. Luxembourg: Publications Office. http://www.cedefop.europa.eu/EN/publications/5069.aspx [accessed 7.5.2012].

Cedefop (2010). *Skills supply and demand in Europe: medium-term forecast up to 2020*. Luxembourg: Publications Office. http://www.cedefop.europa.eu/EN/Files/3052_en.pdf [accessed 7.5.2012].

Cedefop (2011). *What next for skills on the European labour market?* Briefing note; February 2011. http://www.cedefop.europa.eu/EN/Files/9059_en.pdf [accessed 7.5.2012].

Cörvers, F. et al. (2008). *Methodiek arbeidsmarktprognoses en –indicatoren 2007-12 [Methodology of the labour market forecast and indicators 2007-12].* Maastricht: Research Centre for Education and the Labour Market. http://www.roa.unimaas.nl/pdf_publications [accessed 7.5.2012].

Council of the European Union (2007). Council resolution of 15 November 2007 on the new skills for new jobs. *Official journal of the European Union*, C 290, 4.12.2007, Vol. 50, pp. 1-30. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2007:290:0001:0003:EN:PDF [accessed 31.5.2012].

Council of the European Union (2008). *Presidency conclusions*. Brussels European Councils, 13 and 14 March 2008. http://www.consilium.europa.eu/uedocs/cms_Data/docs/pressdata/en/ec/99410.pdf [accessed 31.5.2012].

Doornik, J. (2007). *Ox: an object-oriented matrix language*. London: Timberlake Consultants Press.

Dupuy, A.; Sauermann, J. (2011). *Skills upgrading.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Engle, R.; Granger, C. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica*, Vol. 55, Issue 2, pp. 251–256.

European Commission (2007). *Integrated guidelines for growth and jobs* (2008-10): communication from the Commission to the European Council, COM(2007) 803 final. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52007DC0803(04):EN:HTML [accessed 31.5.2012].

European Commission (2010a). *New skills for new jobs: action now. A report by the expert group on new skills for new jobs prepared for the European Commission*. Luxembourg: Publications Office. http://libserver.cedefop.europa.eu/vetelib/2010/72451.pdf [accessed 7.5.2012].

European Commission (2010b). *Europe 2020: a European strategy for smart, sustainable and inclusive growth: communication from the Commission*. Luxembourg: Publications Office. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF [accessed 7.5.2012].

European Communities (2003). *The European Union labour force survey*. Luxembourg: Publications Office.

Eurostat (1996). *European system of accounts (ESA95)*. European Commission. http://circa.europa.eu/irc/dsis/nfaccount/info/data/ESA95/en/een00sum.htm [accessed 7.5.2012].

Eurostat (2008). *EU labour force survey database: user guide*. European Commission.

Fox, R.; Comerford, B. (2008). Estimating replacement demand: lessons from Ireland. *International journal of manpower,* Vol. 29, Issue 4, pp. 348–361.

Gausas, S. et al. (2011). *Migration*. Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Hecher, D. (2005). Occupational employment projections to 2004-14. *Monthly labor review*, pp. 70–102.

Hendry, A. et al. (1984). Dynamic specification. In: Griliches, Z.; Intriligator, M. (eds). *Handbook of econometrics, Vol. II*. Amsterdam: North Holland.

Hughes, G.; Fox, R. (2005). Country contribution: Ireland*. Feasibility workshop on European skill needs forecasting: information inputs by Member States*, *Pafos, Cyprus, 20 and 21 October 2005*.

Kriechel, B. (2011). *On the flow measure in replacement demand.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Kriechel, B.; Sauermann, J. (2010). *Forecasting skill supply and demand in Europe to 2020: replacement demand.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Kriechel, B.; Wilson, R. (2010). *Forecasting skill supply and demand in Europe: imbalances and Mismatches: reconciling the demand for and supply of skills.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Kriechel, B.; Wilson, R. (2011a). *Qualitative evaluation.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Kriechel, B.; Wilson, R. (2011b). *Interactions and imbalances.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Krusell, P. et al. (2000). Capital-skill complementarity and inequality: a macroeconomic analysis. *Econometrica,* Vol. 68, Issue 5, pp. 1029-1053.

Lazear, E.P.; Shaw, K.L. (eds) (2009). *The structure of wages: an international comparison.* Chicago: University of Chicago Press. NBER Comparative labour markets series.

Lee, K. et al. (1990). Aggregation bias in labour demand equations for the UK economy. In: Barker, T.; Pesaran, M.H. (eds). *Disaggregation in econometric modelling*. Routledge, pp. 113-149.

Livanos, I.; Wilson, R (2007a). Comparing national accounts (NA) and European labour survey (ELFS). Paper presented at the 3rd technical workshop *Medium-term forecast of occupational skills needs in Europe: interim results, Maastricht, 7 and 8 May 2007*, organised by Cedefop .

Livanos, I.; Wilson, R.A. (2007b). The Skillsnet project on medium-term forecasts of occupational skill needs in Europe: expansion demand and changes in occupational and qualifications structure. Paper presented at the 4th technical workshop *Medium-term forecast of occupational skill needs in Europe: final results, Vienna, 8 and 9 November 2007*.

Livanos, I.; Wilson, R.A. (2008). Explaining numbers in stock of those acquiring qualifications at a pan-European level. Paper presented at the technical workshop *Medium-term forecast of occupational skill supply in Europe*,,*Cambridge, 11 and 12 December 2008*, organised by Cedefop.

Livanos, I.; Wilson, R.A. (2010a). *Forecasting skill supply and demand in Europe: modelling the supply of skills.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Livanos, I.; Wilson, R.A (2010b). *Forecasting skill supply and demand for Europe: modelling the demand for skills.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Meagher, G. et al. (2000). *Applied general equilibrium modelling and labour market forecasting.* Centre of policy studies, Monash University.

Meagher, G.A. (1997). Changes in the demand for labour in Australia. *Changing labour markets: prospects for productivity growth*. *Workshop proceedings* /Industry Commission*, Melbourne, 20-21 February 1997.* Melbourne: Industry Commission.

Oxinos, G. et al. (2005). Country contribution: Cyprus. *Feasibility workshop on European skill needs forecasting: information inputs by Member States*, *Pafos, Cyprus, 20 and 21 October 2005*.

Pollitt, H.; Chewpreecha, U. (2008). Development of labour supply projections with the E3ME framework. Paper presented at the technical workshop *Medium-term forecast of occupational skills in Europe: final results, Cambridge, 11 and 12 December 2008*, organised by Cedefop.

Pollitt, H.; Chewpreecha, U. (2009). Macroeconomic scenarios for skill demand and supply projections: the recession and beyond. Paper presented at the Skillsnet technical workshop *Medium-term forecast of skills supply in Europe: interim results*, *Valetta, 15 and 16 October 2009*.

Pollitt, H.; Chewpreecha, U. (2010a). *Links between skills and technological progress in E3ME.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Pollitt, H.; Chewpreecha, U. (2010b). *Macroeconomic scenarios for skill demand and supply projections: the recession and beyond*. Cedefop project on forecasting skill supply and demand in Europe;.technical report. Thessaloniki: Cedefop.

Pollitt, H. et al. (2010). *Forecasting skill supply and demand in Europe to 2020: projections of labour supply and demand from the E3ME model.* Cedefop project on forecasting skills supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Pollit, H et al. (2011). *Forecasting skill supply and demand in Europe to 2020: projections of labour supply and demand from the E3ME model: policy scenarios for year 2.* Cedefop project on forecasting skills supply and demand in Europe. Warwick: IER – Institute for employment research.

Scherer, F.M. (1982). Inter-industry technology flows and productivity growth. *Review of economics and statistics*, Vol. 64, Issue 4, pp. 627-634.

Scherer, F.M. (1984). Using linked patent and R&D data to measure inter-industry technology flows. In: Griliches, Z. (ed.). *R&D, patents and productivity*. Chicago: University of Chicago Press.

Shryock, H.; Siegel, J. (1980). *The methods and materials of demography.* Washington D.C.: Government printing press office.

Spicer, M.; Reade, J. (2005). *Incorporating new EU member economies in multisectoral, multi-region dynamic model: applying shrinkage estimation to E3ME*. COMETR project Competitiveness effects of environmental tax reform. London: Policy Studies Institute.

Stehrer, R.; Ward, T. (2009). Using LFS data for estimating changes in skill demand and supply at a pan-European level. Paper presented at the Skillsnet technical workshop *Medium-term forecast of skills supply in Europe: interim results*, *Valetta, 15 and 16 October 2009*.

Stehrer, R.; Ward, T. (2010). *Forecasting skills supply and demand in Europe: developing a consistent database using LFS data. Final report.* Cedefop project on forecasting skills demand and supply in Europe; technical report. Warwick: IER – Institute for employment research.

Stehrer, R.; Ward, T. (2011). Employment structures by detailed accounts on educational attainment levels, occupation and fields of study. Warwick: IER – Institute for employment research.

Timmer, M. et al. (2007). *EU-KLEMS growth and productivity accounts; Part I: methodology*. http://www.euklems.net/data/EUKLEMS_Growth_and_Productivity_Accounts_Part_I_Methodology.pdf [accessed 31.8.2012].

Willems, E.; de Grip, A. (1993). Forecasting replacement demand by occupation and education. *International journal of forecasting*, Vol. 9, No 2, pp. 173–185.

Wilson, R.A.; Bosworth, D.L. (2006). *Working futures: qualifications report.* Wath on Dearne: Sector Skills Development Agency.

Wilson, R.A.; Pollitt, H. (2011). *Quantitative evaluation of the employment projections.* Cedefop project on forecasting skill supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Wilson, R.A. et al. (2007). *Skills in England 2007: key messages*. Coventry: Learning and Skills Council.

Wilson, R.A. et al. (2010). *Medium-term forecast of skill supply in Europe: overview and synthesis*. Cedefop project on forecasting skill supply and demand in Europe; technical report. Warwick: IER – Institute for employment research.

Yamano, N.; Ahmad, N. (2006). *The OECD input-output database: 2006 edition*. Statistical analysis of science, technology and industry, STI working paper 2006/8.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical association*, Vol. 57, pp. 348–368.

Annex 1

# Acronyms and definitions

| Institutions and organisations | |
|---|---|
| Cedefop | European Centre for the Development of Vocational Training |
| EU | European Union |
| Eurostat | Statistical Office of the European Communities |
| IER | Institute for Employment Research |
| ILO | International Labour Organisation |
| OECD | Organisation for Economic Cooperation and Development |
| ROA | Research Centre for Education and the Labour Market, University of Maastricht |
| Skillsnet | Cedefop's network on early identification of skill needs |
| Unesco | United Nations Educational, Scientific and Cultural Organisation |
| US | United States |

| Others | |
|---|---|
| AMECO | annual macroeconomic database of the European Commission's Directorate-General for Economic and Financial Affairs |
| BALMOD | module to reconcile skill supply and demand projections |
| E3ME | energy-environment-economy model of Europe (multisectoral macroeconomic model) |
| E3ME* | E3ME augmented to include detailed labour supply model |
| EDMOD | module to produce occupational demand projections (expansion demands) |
| ESA95 | European system of accounts |
| EU-27 | European Union of 27 Member States |
| EU-27+ | European Union of 27 Member States plus Norway and Switzerland |
| EU-LFS | European Union labour force survey |
| FlowMOD | module of flows into and out from the education system |
| GDP | gross domestic product |
| ISCED | International standard classification of education |
| ISCO | International standard classification of occupations |
| LFS | labour force survey |
| LMAR | labour market accounts residuals |

| | |
|---|---|
| NA | national accounts |
| p.a. | per annum |
| QMOD | module to produce qualification projections |
| RDMOD | mdule to produce projections of replacement demands |
| StockMOD | module of numbers acquiring qualifications (stocks) |

**Definitions of terms used**

| | |
|---|---|
| conceptual framework | The general theoretical and methodological approach to modelling and projecting the demand for and supply of skills. |
| employment | The number of people in work (headcount), national accounts definition, (or the number of jobs in some cases), split by various dimensions, including sector, occupation, gender and highest qualification held. |
| labour force | The number of people economically active (the sum over the various age ranges of the working age population * the relevant labour market participation rate) which includes employed and unemployed. |
| population (15+) | Anyone of age 15 or over is classified as part of the population in the context of the model. People over 65 are included in this definition, as these age groups have participation rates greater than zero. |
| working age population | Anyone of age 15-64 is classified as part of the working age population. |
| participation or activity rate | The percentage of the population that is either employed or unemployed (ILO definition of labour force). This is differentiated by gender and age group. |
| qualifications | This term refers to the highest level of education/qualification held by the individual. The ISCED classification is used for this purpose. The most aggregate level distinguishes three main levels of education/qualification: high (ISCED 5-6), medium (ISCED 3-4, excluding 3c short) and low (ISCED 0-2, plus 3c short). |
| demand | In the context of the model, labour demand is taken to be the same as employment levels (number of jobs available). It does not include (for example) unfilled vacancies. |
| supply | In the context of the model, labour supply is taken to be the same as the labour force. |

Annex 2
# Classifications and aggregations used in the modelling framework

## Industries and sectors

Table A1. **Aggregation of NACE Rev 1.1 two and three digit industries to 41 industries**

| 41-industry [NACE] | | NACE Rev 1.1 [NACE] |
|---|---|---|
| 1 | Agriculture, etc. [01-05] | Agriculture, hunting and related service activities [01] |
| | | Forestry, logging and related service activities [02] |
| | | Fishing, fish farming and related service activities [05] |
| 2 | Coal [10] | Mining of coal and lignite; extraction of peat [10] |
| 3 | Oil and gas, etc. [11, 12] | Extraction of crude petroleum and natural gas; service activities incidental to oil and gas extraction, excluding surveying [11] |
| | | Mining of uranium and thorium ores [12] |
| 4 | Other mining [13, 14] | Mining of metal ores [13] |
| | | Other mining and quarrying [14] |
| 5 | Food, drink and tobacco [15, 16] | Manufacture of food products and beverages [15] |
| | | Manufacture of tobacco products [16] |
| 6 | Textiles, clothing and leather [17-19] | Manufacture of textiles [17] |
| | | Manufacture of wearing apparel; dressing and dyeing of fur [18] |
| | | Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear [19] |
| 7 | Wood and paper [20, 21] | Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials [20] |
| | | Manufacture of pulp, paper and paper products [21] |
| 8 | Printing and publishing [22] | Publishing, printing and reproduction of recorded media [22] |
| 9 | Manufactured fuels [23] | Manufacture of coke, refined petroleum products and nuclear fuel [23] |
| 10 | Pharmaceuticals [24.4] | Manufacture of pharmaceuticals, medicinal chemicals and botanical products [24.4] |

| | | |
|---|---|---|
| 11 | Chemicals n.e.s. [24(ex24.4)] | Manufacture of chemicals and chemical products (except pharmaceuticals, etc.) [24 (ex 24.4)] |
| 12 | Rubber and plastics [25] | Manufacture of rubber and plastic products [25] |
| 13 | Non-metallic mineral products [26] | Manufacture of other non-metallic mineral products [26] |
| 14 | Basic metals [27] | Manufacture of basic metals [27] |
| 15 | Metal goods [28] | Manufacture of fabricated metal products, except machinery and equipment [28] |
| 16 | Mechanical engineering [29] | Manufacture of machinery and equipment n.e.c. [29] |
| 17 | Electronics [30, 32] | Manufacture of office machinery and computers [30] |
| | | Manufacture of radio, television and communication equipment and apparatus [32] |
| 18 | Electrical engineering and instruments [31, 33] | Manufacture of electrical machinery and apparatus n.e.c. [31] |
| | | Manufacture of medical, precision and optical instruments, watches and clocks [33] |
| 19 | Motor vehicles [34] | Manufacture of motor vehicles, trailers and semi-trailers [34] |
| 20 | Other transport equipment [35] | Manufacture of other transport equipment [35] |
| 21 | Manufacturing nes [36, 37] | Manufacture of furniture; manufacturing n.e.c. [36] |
| | | Recycling [37] |
| 22 | Electricity [40.1, 40.3] | Electricity, steam and hot water supply [40.1, 40.3] |
| 23 | Gas supply [40.2] | Manufacture of gas; distribution of gaseous fuels through mains [40.2] |
| 24 | Water supply [41] | Collection, purification and distribution of water [41] |
| 25 | Construction [45] | Construction [45] |
| 26 | Distribution [50, 51] | Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel [50] |
| | | Wholesale trade and commission trade, except of motor vehicles and motorcycles [51] |
| 27 | Retailing [52] | Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods [52] |
| 28 | Hotels and catering [55] | Hotels and restaurants [55] |
| 29 | Land transport, etc.[60, 63] | Land transport; transport via pipelines [60] |
| | | Supporting and auxiliary transport activities; activities of travel agencies [63] |
| 30 | Water transport [61] | Water transport [61] |

| | | |
|---|---|---|
| 31 | Air transport [62] | Air transport [62] |
| 32 | Communications [64] | Post and telecommunications [64] |
| 33 | Banking and finance [65, 67] | Financial intermediation, except insurance and pension funding [65] |
| | | Activities auxiliary to financial intermediation [67] |
| 34 | Insurance [66] | Insurance and pension funding, except compulsory social security [66] |
| 35 | Computing services [72] | Computer and related activities [72] |
| 36 | Professional services [70, 71, 73, 74.1-74.4] | Real estate activities [70] |
| | | Renting of machinery and equipment without operator and of personal and household goods [71] |
| | | Research and development [73] |
| | | Other business activities (professional services) [74.1-74.4] |
| 37 | Other Business services [74.5-74.8] | Other business activities (business services) [74.5-74.8] |
| 38 | Public administration and Defence [75] | Public administration and defence; compulsory social security [75] |
| 39 | Education [80] | Education [80] |
| 40 | Health and social work [85] | Health and social work [85] |
| 41 | Miscellaneous services [90-93, 95-97, 99] | Sewage and refuse disposal, sanitation and similar activities [90] |
| | | Activities of membership organisations n.e.c. [91] |
| | | Recreational, cultural and sporting activities [92] |
| | | Other service activities [93] |
| | | Activities of households as employers of domestic staff [95] |
| | | Undifferentiated goods producing activities of private households for own use [96] |
| | | Undifferentiated services producing activities of private households for own use [97] |
| | | Extra-territorial organisations and bodies [99] |

(a) *Source*: http://forum.europa.eu.int/irc/dsis/employment/info/data/eu_lfs/Related_documents/Nace_Rev_1.1.htm

Table A2. **Aggregation of 41-industry to 6-industry**

| 6-industry [NACE] | | 41-industry [NACE] | |
|---|---|---|---|
| 1 | Primary sector and utilities [01-14, 40, 41] | 1 | Agriculture, etc.[01-05] |
| | | 2 | Coal [10] |
| | | 3 | Oil and gas, etc.[11, 12] |
| | | 4 | Other mining [13, 14] |
| | | 22 | Electricity [40.1, 40.3] |
| | | 23 | Gas supply [40.2] |
| | | 24 | Water supply [41] |
| 2 | Manufacturing [15-37] | 5 | Food, drink and tobacco [15, 16] |
| | | 6 | Textiles, clothing and leather [17-19] |
| | | 7 | Wood and paper [20, 21] |
| | | 8 | Printing and publishing [22] |
| | | 9 | Manufactured fuels [23] |
| | | 10 | Pharmaceuticals [24.4] |
| | | 11 | Chemicals nes [24(ex24.4)] |
| | | 12 | Rubber and plastics [25] |
| | | 13 | Non-metallic mineral products [26] |
| | | 14 | Basic metals [27] |
| | | 15 | Metal goods [28] |
| | | 16 | Mechanical engineering [29] |
| | | 17 | Electronics [30, 32] |
| | | 18 | Electrical engineering and instruments [31, 33] |
| | | 19 | Motor vehicles [34] |
| | | 20 | Other transport equipment [35] |
| | | 21 | Manufacturing nes [36, 37] |
| 3 | Construction [45] | 25 | Construction [45] |
| 4 | Distribution and transport [50-64] | 26 | Distribution [50, 51] |
| | | 27 | Retailing [52] |
| | | 28 | Hotels and catering [55] |
| | | 29 | Land transport, etc.[60, 63] |
| | | 30 | Water transport [61] |
| | | 31 | Air transport [62] |
| | | 32 | Communications [64] |
| 5 | Business and other services [65-74, 90-99] | 33 | Banking and finance [65, 67] |
| | | 34 | Insurance [66] |
| | | 35 | Computing services [72] |
| | | 36 | Professional services [70, 71, 73, 74.1-74.4] |
| | | 37 | Other Business services [74.5-74.8] |
| | | 41 | Miscellaneous services [90-93, 95, 99] |
| 6 | Non-marketed services [75, 80, 85] | 38 | Public administration and defence [75] |
| | | 39 | Education [80] |
| | | 40 | Health and social work [85] |

# Occupations

Table A3. **ISCO**

| | |
|---|---|
| **Major group 1: legislators, senior officials and managers** | |
| 11 | Legislators and senior officials |
| 12 | Corporate managers |
| 13 | Managers of small enterprises |
| **Major group 2: professionals** | |
| 21 | Physical, mathematical and engineering science professionals |
| 22 | Life science and health professionals |
| 23 | Teaching professionals |
| 24 | Other professionals |
| **Major group 3: technicians and associate professionals** | |
| 31 | Physical and engineering science associate professionals |
| 32 | Life science and health associate professionals |
| 33 | Teaching associate professionals |
| 34 | Other associate professionals |
| **Major group 4: clerks** | |
| 41 | Office clerks |
| 42 | Customer services clerks |
| **Major group 5: service workers and shop and market sales workers** | |
| 51 | Personal and protective services workers |
| 52 | Models, salespersons and demonstrators |
| **Major group 6: Skilled agricultural and fishery workers** | |
| 61 | Skilled agricultural and fishery workers |
| **Major group 7: craft and related trades workers** | |
| 71 | Extraction and building trades workers |
| 72 | Metal, machinery and related trades workers |
| 73 | Precision, handicraft, craft printing and related trades workers |
| 74 | Other craft and related trades workers |
| **Major group 8: plant and machine operators and assemblers** | |
| 81 | Stationary plant and related operators |
| 82 | Machine operators and assemblers |
| 83 | Drivers and mobile plant operators |
| **Major group 9: elementary occupations** | |
| 91 | Sales and services elementary occupations |
| 92 | Agricultural, fishery and related labourers |
| 93 | Labourers in mining, construction, manufacturing and transport |
| **Major group 0: armed forces** | |

## Qualifications

| Level of qualification | |
|---|---|
| Low | (Pre)primary and lower secondary (ISCED 0-2) |
| Medium | Upper and post-secondary (ISCED 3-4) |
| High | Tertiary (ISCED 5-6) |

**ISCED 0: pre-primary education**

Programmes at level 0, (pre-primary) defined as the initial stage of organised instruction, are designed primarily to introduce young children to a school-type environment, i.e. to provide a bridge between the home and a school-based atmosphere. Upon completion of these programmes, children continue their education at level 1 (primary education).

**ISCED 1: primary education or first stage of basic education**

Programmes at level 1 are normally designed on a unit or project basis to give students a sound basic education in reading, writing and mathematics along with an elementary understanding of other subjects such as history, geography, natural science, social science, art and music. In some cases religious instruction is featured. The core at this level consists of education provided for children, the customary or legal age of entrance being not younger than five years or older than seven years. This level covers, in principle, six years of full-time schooling.

**ISCED 2: lower secondary education or second stage of basic education**

The contents of education at this stage are typically designed to complete the provision of basic education which began at ISCED level 1. In many, if not most countries, the educational aim is to lay the foundation for lifelong learning and human development. Programmes are usually on a more subject oriented pattern using more specialised teachers and more often several teachers conducting classes in their field of specialisation. The full implementation of basic skills occurs. The end of this level often coincides with the end of compulsory schooling where it exists.

**ISCED 3: upper secondary education**

This level of education typically begins at the end of full-time compulsory education for those countries that have a system of compulsory education. More specialisation may be observed than at ISCED 2 and often teachers need to be more qualified or specialised than for ISCED 2. The entrance age is typically 15 to 16 years. The educational programmes included typically require the completion of some nine years of full-time education (since the beginning of level

1) for admission or a combination of education and vocational or technical experience.

ISCED 3A: programmes designed to provide direct access to ISCED 5A;

ISCED 3B: programmes designed to provide direct access to ISCED 5B;

ISCED 3C: programmes not designed to lead to ISCED 5A or 5B.

**ISCED 4: post-secondary non tertiary education**

ISCED 4 captures programmes that straddle the boundary between upper secondary and post-secondary education from an international point of view, even though they might clearly be considered as upper secondary or post-secondary programmes in a national context. Considering their content, these programmes cannot be regarded as tertiary programmes. They are often not significantly more advanced than programmes at ISCED 3 but they serve to broaden the knowledge of participants who have already completed a programme at level 3.

Typical examples are programmes designed to prepare students for studies at level 5 who, although having completed ISCED 3, did not follow a curriculum which would allow entry to level 5, i.e. pre-degree foundation courses or short vocational programmes. Second cycle programmes can be included as well.

ISCED 4A: see text for ISCED 3

ISCED 4B: see text for ISCED 3

ISCED 4C: see text for ISCED 3

**ISCED 5: first stage of tertiary education (not leading directly to an advanced research qualification)**

This level consists of tertiary programmes having an educational content more advanced than those offered at levels 3 and 4. Entry to these programmes normally requires the successful completion of ISCED 3A or 3B or a similar qualification at ISCED 4A. They do not lead to the award of an advanced research qualification (ISCED 6). These programmes must have a cumulative duration of at least two years.

ISCED 5A: programmes that are largely theoretically based and are intended to provide sufficient qualifications for gaining entry into advanced research programmes and professions with high skills requirements.

ISCED 5B: programmes that are practically oriented/occupationally specific and are mainly designed for participants to acquire the practical skills and know-how needed for employment in a particular occupation or trade or class of occupations or trades, the successful completion of which usually provides the participants with a labour-market relevant qualification.

**ISCED 6: second stage of tertiary education (leading to an advanced research qualification)**

This level is reserved for tertiary programmes which lead to the award of an advanced research qualification. The programmes, therefore, are devoted to advanced study and original research and not based on course-work only. They typically require the submission of a thesis or dissertation of publishable quality which is the product of original research and represents a significant contribution to knowledge. They prepare graduates for faculty posts in institutions offering ISCED 5A programmes, as well as research posts in government, industry, etc.

Documentation by EULFS: Levels of education and training ISCED 1997 (http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/Related_documents/ISCED_EN.htm)

Table A4. **Division of education by fields of study**

| | |
|---|---|
| 01 | Agriculture and veterinary |
| 02 | Computer science |
| 03 | Engineering, manufacturing and construction |
| 04 | Foreign languages |
| 05 | General programmes |
| 06 | Health and welfare |
| 07 | Humanities, languages and arts |
| 08 | Life science (including Biology and Environmental science) |
| 09 | Mathematics and statistics |
| 10 | Physical science (including Physics, Chemistry and Earth science) |
| 11 | Services |
| 12 | Social sciences, business and law |
| 13 | Teacher training and education science |
| 14 | Computer use |
| 15 | Science, mathematics and computing |

# Annex 3
# Distribution of fields of study across the occupations

| ISCO | Fields of study | BG | CZ | DE | ES | IT | UK | EU-27 |
|---|---|---|---|---|---|---|---|---|
| 131 | Agriculture and veterinary | 3.7 | 5.6 | 2.0 | 0.7 | 1.7 | 2.9 | 3.1 |
| | Computer science | 0.4 | 0.3 | 0.3 | 1.0 | 0.8 | 1.1 | 0.7 |
| | Engineering, manufacturing and construction | 49.9 | 46.5 | 26.4 | 7.6 | 13.7 | 21.4 | 18.7 |
| | Foreign languages | 0.4 | 0.2 | 0.3 | 0.0 | 1.1 | 1.0 | 0.6 |
| | General programmes | 13.6 | 4.4 | 4.9 | 17.4 | 2.9 | 0.5 | 8.9 |
| | Health and welfare | 2.0 | 2.0 | 2.8 | 2.1 | 0.5 | 4.4 | 2.2 |
| | Humanities, languages and arts | 2.4 | 2.4 | 2.7 | 1.7 | 3.5 | 6.6 | 3.4 |
| | Life science (including Biology and Environmental science) | 0.1 | 0.5 | 0.2 | 0.2 | 0.3 | 1.6 | 0.4 |
| | Mathematics and statistics | 0.4 | 0.2 | 0.2 | 0.1 | 0.1 | 0.4 | 0.2 |
| | Physical science (including Physics, Chemistry and Earth science) | 0.6 | 1.0 | 0.6 | 0.7 | 0.1 | 1.9 | 0.7 |
| | Services | 5.0 | 9.6 | 9.8 | 1.9 | 3.1 | 8.1 | 5.5 |
| | Social sciences, business and law | 15.4 | 20.8 | 32.3 | 12.4 | 19.4 | 20.1 | 19.6 |
| | Teacher training and education science | 2.8 | 5.0 | 2.3 | 1.4 | 2.6 | 2.1 | 2.0 |
| | Computer use | 0.1 | 0.0 | 0.4 | 0.0 | 0.1 | 0.6 | 0.2 |
| | Science, mathematics and computing | 0.0 | 0.0 | 0.0 | 0.6 | 3.9 | 0.7 | 1.4 |
| | NA | 3.2 | 1.5 | 14.7 | 52.1 | 46.4 | 26.5 | 32.3 |
| 241 | Agriculture and veterinary | 0.9 | 2.8 | 1.0 | 0.3 | 0.8 | 0.4 | 1.3 |
| | Computer science | 1.1 | 0.5 | 1.8 | 0.4 | 0.9 | 1.9 | 1.5 |
| | Engineering, manufacturing and construction | 11.8 | 17.3 | 14.6 | 3.2 | 11.2 | 5.0 | 10.1 |
| | Foreign languages | 0.9 | 0.2 | 1.0 | 0.0 | 2.0 | 2.2 | 1.2 |
| | General programmes | 1.6 | 6.1 | 4.5 | 1.5 | 2.9 | 0.3 | 4.1 |
| | Health and welfare | 0.8 | 1.7 | 2.0 | 2.2 | 0.3 | 2.6 | 2.3 |
| | Humanities, languages and arts | 1.0 | 1.9 | 2.9 | 5.0 | 3.3 | 5.9 | 3.7 |
| | Life science (including Biology and Environmental science) | 0.4 | 0.2 | 0.7 | 0.9 | 0.5 | 1.6 | 0.9 |
| | Mathematics and statistics | 0.3 | 1.1 | 0.7 | 1.7 | 0.5 | 4.1 | 1.5 |
| | Physical science (including Physics, Chemistry and Earth science) | 0.8 | 1.8 | 2.2 | 1.0 | 0.7 | 3.3 | 1.5 |
| | Services | 1.4 | 3.1 | 2.2 | 2.5 | 0.7 | 1.5 | 2.3 |
| | Social sciences, business and law | 77.5 | 59.6 | 58.4 | 77.3 | 67.6 | 60.9 | 62.1 |
| | Teacher training and education science | 1.2 | 3.1 | 3.3 | 3.0 | 1.5 | 1.8 | 2.8 |
| | Computer use | 0.2 | 0.1 | 0.8 | 0.0 | 0.2 | 0.4 | 0.4 |
| | Science, mathematics and computing | 0.0 | 0.0 | 0.0 | 1.0 | 3.1 | 2.6 | 1.2 |
| | NA | 0.1 | 0.3 | 3.7 | 0.2 | 3.8 | 5.4 | 3.1 |
| 343 | Agriculture and veterinary | 1.7 | 5.5 | 0.6 | 0.4 | 0.8 | 1.0 | 1.4 |
| | Computer science | 0.6 | 0.1 | 0.4 | 1.8 | 1.3 | 1.7 | 0.9 |
| | Engineering, manufacturing and construction | 12.7 | 13.5 | 8.3 | 2.7 | 7.2 | 2.3 | 7.2 |
| | Foreign languages | 0.4 | 0.2 | 1.3 | 0.0 | 2.5 | 2.0 | 1.3 |
| | General programmes | 12.3 | 10.8 | 3.7 | 26.0 | 4.0 | 0.8 | 9.4 |
| | Health and welfare | 0.8 | 2.4 | 2.6 | 1.7 | 0.6 | 2.0 | 2.1 |
| | Humanities, languages and arts | 1.6 | 1.3 | 1.2 | 4.8 | 4.5 | 4.4 | 3.2 |

| ISCO | Fields of study | BG | CZ | DE | ES | IT | UK | EU-27 |
|---|---|---|---|---|---|---|---|---|
| | Life science (including Biology and Environmental science) | 0.5 | 0.2 | 0.1 | 0.4 | 0.2 | 1.2 | 0.3 |
| | Mathematics and statistics | 0.7 | 0.2 | 0.2 | 0.4 | 0.3 | 1.5 | 0.4 |
| | Physical science (including Physics, Chemistry and Earth science) | 0.4 | 0.7 | 0.3 | 0.6 | 0.2 | 1.9 | 0.5 |
| | Services | 2.7 | 4.3 | 3.5 | 1.3 | 1.9 | 2.3 | 3.0 |
| | Social sciences, business and law | 62.9 | 58.5 | 68.6 | 44.9 | 59.5 | 62.9 | 59.3 |
| | Teacher training and education science | 2.2 | 1.6 | 1.8 | 3.4 | 4.7 | 1.6 | 2.5 |
| | Computer use | 0.0 | 0.1 | 0.6 | 0.1 | 0.3 | 1.8 | 0.3 |
| | Science, mathematics and computing | 0.0 | 0.0 | 0.0 | 1.3 | 3.8 | 1.9 | 1.3 |
| | NA | 0.5 | 0.6 | 6.7 | 10.3 | 8.0 | 10.7 | 7.0 |
| 712 | Agriculture and veterinary | 3.0 | 1.5 | 0.9 | 0.4 | 1.3 | 0.7 | 1.4 |
| | Computer science | 0.2 | 0.0 | 0.1 | 0.6 | 0.3 | 0.7 | 0.3 |
| | Engineering, manufacturing and construction | 42.6 | 89.3 | 78.5 | 9.3 | 11.6 | 65.2 | 40.7 |
| | Foreign languages | 0.1 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.1 |
| | General programmes | 14.4 | 0.6 | 0.9 | 10.3 | 0.5 | 0.9 | 5.1 |
| | Health and welfare | 0.0 | 0.1 | 0.3 | 0.4 | 0.0 | 0.4 | 0.3 |
| | Humanities, languages and arts | 0.2 | 0.4 | 0.5 | 0.4 | 0.7 | 1.9 | 0.8 |
| | Life science (including Biology and Environmental science) | 0.2 | 0.0 | 0.0 | 0.1 | 0.0 | 0.3 | 0.1 |
| | Mathematics and statistics | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 | 0.4 | 0.1 |
| | Physical science (including Physics, Chemistry and Earth science) | 0.0 | 0.1 | 0.1 | 0.3 | 0.1 | 0.4 | 0.2 |
| | Services | 2.1 | 2.2 | 1.1 | 0.3 | 1.0 | 1.5 | 1.7 |
| | Social sciences, business and law | 0.8 | 0.8 | 0.9 | 1.6 | 2.5 | 2.9 | 1.9 |
| | Teacher training and education science | 0.2 | 0.0 | 0.1 | 0.2 | 0.2 | 0.4 | 0.3 |
| | Computer use | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.1 |
| | Science, mathematics and computing | 0.0 | 0.0 | 0.0 | 0.1 | 1.1 | 0.3 | 0.3 |
| | NA | 36.1 | 5.0 | 16.6 | 76.1 | 80.3 | 23.8 | 46.7 |
| 832 | Agriculture and veterinary | 4.3 | 5.6 | 4.6 | 0.5 | 1.8 | 1.9 | 3.3 |
| | Computer science | 0.1 | 0.0 | 0.2 | 0.5 | 0.6 | 1.1 | 0.3 |
| | Engineering, manufacturing and construction | 57.7 | 77.0 | 50.3 | 10.2 | 14.4 | 27.0 | 34.9 |
| | Foreign languages | 0.0 | 0.0 | 0.1 | 0.0 | 0.1 | 0.4 | 0.1 |
| | General programmes | 14.2 | 2.2 | 2.4 | 13.6 | 0.9 | 0.8 | 6.6 |
| | Health and welfare | 0.3 | 0.3 | 0.7 | 0.6 | 0.2 | 1.7 | 0.6 |
| | Humanities, languages and arts | 0.4 | 0.4 | 0.9 | 0.5 | 1.2 | 3.8 | 1.1 |
| | Life science (including Biology and Environmental science) | 0.1 | 0.0 | 0.2 | 0.0 | 0.0 | 0.7 | 0.1 |
| | Mathematics and statistics | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.3 | 0.1 |
| | Physical science (including Physics, Chemistry and Earth science) | 0.1 | 0.3 | 0.2 | 0.5 | 0.1 | 0.6 | 0.3 |
| | Services | 6.9 | 5.2 | 13.1 | 0.6 | 1.4 | 12.2 | 7.5 |
| | Social sciences,business and law | 1.0 | 2.3 | 6.8 | 3.9 | 5.5 | 9.8 | 5.1 |
| | Teacher training and education science | 0.5 | 0.0 | 0.6 | 0.3 | 0.4 | 0.7 | 0.4 |
| | Computer use | 0.0 | 0.1 | 0.1 | 0.0 | 0.1 | 1.1 | 0.1 |
| | Science, mathematics and computing | 0.0 | 0.0 | 0.0 | 0.3 | 1.2 | 0.4 | 0.4 |
| | NA | 14.5 | 6.6 | 19.7 | 68.4 | 72.0 | 37.5 | 38.9 |

# Annex 4
# Summary of SORT algorithm functioning

The sorting algorithm at the heart of BALMOD is designed to reconcile the projections from the stock model of supply (numbers available by the three qualification levels) with those from the demand for qualifications model (number of jobs requiring particular qualification levels). The former provides a view of supply-side developments (the overall numbers of people who have acquired qualifications at the three different levels who are actively searching for work), while the latter is more concerned with changing demand for qualifications within occupations (the number of jobs available requiring particular levels of qualifications).

The module also has to deal with differences between the various estimates of employment used in E3ME (based on national accounts and LFS data) and the labour-market accounts residual (LMAR), which arises in part because of such discrepancies but which is also affected by other issues, including measurement error. The main employment measure used in E3ME is a national accounts-based one. This is referred to as unconstrained estimates of employment. All the estimates by sector and occupation are based on this. A second measure, based on LFS information and Eurostat demographic data, is implicit in the modelling of labour supply. This is referred to as supply in employment. The two differ for various reasons, encompassed under the heading of the LMAR. These include:

- double jobbing (some have more than one job);
- distinction between residence and workplace (many people do not live in the same country as they work; this is especially significant for some small countries such as Luxembourg);
- government training and related schemes (which may count as being in the labour force but not as being in employment);
- different definitions of unemployment (ILO versus measures of claimants to benefits);
- statistical errors (in measures of employment, unemployment and related indicators, including sampling and measurements errors);
- other differences due to use of different data sources; treatment of the armed forces and nationals working abroad.

The sorting model uses an iterative RAS procedure to reconcile two sets of estimates of employment, changing the overall qualification shares from the demand for qualifications model (QUALMOD) to match those from the stock model of supply (STOCKMOD). This is done while at the same time maintaining the patterns of occupational deployment and ensuring a plausible pattern of unemployment rates for the different qualification categories. It therefore focuses on the occupations people with different qualifications end up in.

Overall unemployment levels are taken from E3ME. This is taken as exogenous for these purposes. The overall level of unemployment is shared out among qualification categories, based on an extrapolation of patterns from historical LFS data. In the current versions it is assumed that the relative rates of unemployment for the three broad qualification categories are maintained. Checks are made to see that this results in plausible unemployment levels for the three qualification categories. The implied unemployment levels by qualification are then deducted from the overall supply numbers to get the numbers of people in employment by qualification level (supply in employment). The sorting model then reconciles these estimates with the number of jobs available (unconstrained estimates). This is done by altering the shares of people with the three different qualification levels employed within each occupation, until the overall numbers match the numbers of people available.

The final results may provide indications of overqualification or underqualification of people in different occupations, depending on the overall demand-supply balance.

The constraint (matching of numbers by the three qualifications levels) is imposed at the two-digit occupational level. The key dimensions in the SORT routine are:

- occupation (27);
- qualification level (3);
- sector (41).

(note: the results in several summary tables in the imbalances workbooks where this process is undertaken show outcomes for aggregate one digit occupational groups and six broad sectors only).

The sorting model operates for each country separately. There are assumed to be no adjustments through cross-border flows (migration or commuting).

There is then one final step in which the final outcomes from the sort routine are scaled to match the original E3ME employment totals, to deal with the LMAR discrepancy.

# Annex 5
# Questionnaire for evaluation 'Future skill needs in Europe'

The following short questionnaire will deal with several aspects of the medium-term skill needs forecast from Cedefop. Please help us improve the forecast by giving us your opinion. The results will be presented and reported in aggregated form.

The results of the forecast were published in in different formats. There were concise summaries of results in briefing notes, while the full results can be found in the (synthesis) reports. Background papers give some of the technical details underlying the forecast.

Which reports are you aware of and how useful were they?

| | Not received/ Don't know this | Received, but haven't read it | Received, but found it not interesting | Received, and found it somewhat interesting | Received, and found it very interesting |
|---|---|---|---|---|---|
| Briefing notes | [ ] | [ ] | [ ] | [ ] | [ ] |
| (Synthesis) report | [ ] | [ ] | [ ] | [ ] | [ ] |
| Background technical reports | [ ] | [ ] | [ ] | [ ] | [ ] |

Are there national skills forecast in your own country?

| | |
|---|---|
| [ ] | yes, more detailed |
| [ ] | yes, about the same level of detail |
| [ ] | yes, less detailed |
| [ ] | no (skip to 'Information text') |

What is the time frame of the skill forecast in your country?

| | |
|---|---|
| [ ] | less than 2 years ahead |
| [ ] | 2-5 years |
| [ ] | 6-10 years |
| [ ] | more than 10 years |

### The following questions are about the Cedefop skills forecast to 2020

Which part of the skills forecast have you been using? *(tick all applicable)*

| | |
|---|---|
| [ ] | general results/summary |
| [ ] | workbooks |
| [ ] | methodology reports/background reports |
| [ ] | none (skip to 'Information text') |

What did you use the forecast for?

| | |
|---|---|
| [ ] | to make policy recommendations |
| [ ] | for cross country comparisons |
| [ ] | for further research |
| [ ] | methodology adoption and implementation |

Please rate the following aspects of the model in terms of importance to you

| | (1) Very important | (2) Important | (3) Neutral | (4) Less important | (5) Not important |
|---|---|---|---|---|---|
| Forecast on number of workers in sectors | [ ] | [ ] | [ ] | [ ] | [ ] |
| Forecast on demand for workers in occupations | [ ] | [ ] | [ ] | [ ] | [ ] |
| Forecast on demand for workers by education level | [ ] | [ ] | [ ] | [ ] | [ ] |
| Forecast on labour supply by education | [ ] | [ ] | [ ] | [ ] | [ ] |
| Forecast on imbalances | [ ] | [ ] | [ ] | [ ] | [ ] |

**Please rate the outcomes of the Skillsnet forecast in terms of its plausibility for your country and for pan−European comparison**

| | (1) plausible | (2) partly plausible | (3) neutral | (4) partly implausible | (5) im-plausible |
|---|---|---|---|---|---|
| How plausible are the outcomes of the sectoral level forecasts? | | | | | |
| On a national level | [ ] | [ ] | [ ] | [ ] | [ ] |
| On a pan-European level | [ ] | [ ] | [ ] | [ ] | [ ] |
| How plausible are the outcomes of the occupation projections (expansion demands)? | | | | | |
| On a national level | [ ] | [ ] | [ ] | [ ] | [ ] |
| On a pan-European level | [ ] | [ ] | [ ] | [ ] | [ ] |
| How plausible are the replacement demand projections? | | | | | |
| On a national level | [ ] | [ ] | [ ] | [ ] | [ ] |
| On a pan-European level | [ ] | [ ] | [ ] | [ ] | [ ] |
| How plausible are the outcomes of the demand for qualification level? | | | | | |
| On a national level | [ ] | [ ] | [ ] | [ ] | [ ] |
| On a pan-European level | [ ] | [ ] | [ ] | [ ] | [ ] |
| How plausible are the outcomes of the supply forecast by qualification level? | | | | | |
| On a national level | [ ] | [ ] | [ ] | [ ] | [ ] |
| On a pan-European level | [ ] | [ ] | [ ] | [ ] | [ ] |
| How plausible are the outcomes of the demand for qualification level? | | | | | |
| On a national level | [ ] | [ ] | [ ] | [ ] | [ ] |
| On a pan-European level | [ ] | [ ] | [ ] | [ ] | [ ] |

Overall, how do the outcomes of the Cedefop forecast relate to national forecasts and/or national expectations for the future?

| | |
|---|---|
| [ ] | Very close to national forecast/expectations |
| [ ] | Same direction of outcomes |
| [ ] | Partially contradicting outcomes |
| [ ] | Contradicts national forecasts/expectations |
| [ ] | Don't know |

How important is the availability of more detailed results and underlying data in the form of workbooks to you?

| (1) Very important | (2) Important | (3) Neutral | (4) Less important | (5) Not important |
|---|---|---|---|---|
| [ ] | [ ] | [ ] | [ ] | [ ] |

How important is it to allow the user to change underlying data/forecast assumptions in the workbook?

| (1) Very important | (2) Important | (3) Neutral | (4) Less important | (5) Not important |
|---|---|---|---|---|
| [ ] | [ ] | [ ] | [ ] | [ ] |

In which direction should the project be developed in the future? (Please choose up to three categories)

| | |
|---|---|
| [ ] | Policy simulations |
| [ ] | Technological/organisational changes |
| [ ] | Simulations individual country scenarios |
| [ ] | Individual sector scenarios more detailed |
| [ ] | skills/competences forecast |
| [ ] | Other (please specify): |

How much value added is there by the pan-European projections compared with similar efforts already available at national level?

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

Thank you for participating in the survey.

If you have any questions or comments, please feel free to contact:

Ben.Kriechel@maastrichtuniversity.nl

or

Vladimir.Kvetan@cedefop.europa.eu

European Centre for the Development of Vocational Training

# Skills supply and demand in Europe

## Methodological framework

Luxembourg:
Publications Office of the European Union

2012 – VI, 144 p. – 21 x 29.7 cm

ISBN 978-92-896-1112-1
ISSN 1831-5860
doi: 10.2801/85871
Cat. No: TI-BC-12-006-EN-N

Free download at:
http://www.cedefop.europa.eu/EN/Files/5525_en.pdf

Free of charge – 5525 EN –

# Skills supply and demand in Europe

## Methodological framework

This publication provides an overview of the methodology underpinning the Cedefop skills supply and demand forecast. The aim is to share the knowledge acquired during the development of the different systems, modules and models and to highlight the results, which have already aroused considerable interest. The structure and robustness of the methods used increases the clarity and reliability of the results.

Cedefop's forecast is not intended to replace forecasting efforts in individual countries. However, sharing the knowledge and experience from the European level can help to improve the methods used in those countries and to resolve issues that may be faced there. At the same time, Cedefop's forecast can inspire new forecasting initiatives and initiate further discussion. The feedback received can help to improve Cedefop's methods and make the European forecast even more precise and reliable.

TI-BC-12-006-EN-N

Europe 123, 570 01 Thessaloniki (Pylea), GREECE
Postal address: PO Box 22427, 551 02 Thessaloniki, GREECE
Tel. +30 2310490111, Fax +30 2310490020
E-mail: info@cedefop.europa.eu

visit our portal **www.cedefop.europa.eu**

**5525 EN**

http://www.cedefop.europa.eu/EN/Files/5525_en.pdf

Publications Office

ISBN 978-92-896-1112-1