CANADIAN COUNCIL ON LEARNING CCL CCA CONSEIL CANADIEN SUR L'APPRENTISSAGE

Is it Possible to Accurately Forecast Labour Market Needs?

Prepared by

Canadian Council on Learning

For

British Columbia Ministry of Advanced Education

January 10, 2007

EXECUTIVE SUMMARY

Since its inception, labour market forecasting has been the subject of much debate among economists. The poor track record of forecasts and the complex task of "tuning" the labour market (Heijke, 1996) have led to the view that labour market planning is neither necessary nor useful (Borghans & Willems, 1998). The purpose of this review is to systematically and transparently gather, analyze and synthesize research devoted to discussing and determining whether it is possible to accurately predict labour market needs.

There are two major approaches to occupational forecasting: workforce projection and labour market analysis (signalling). Workforce projection produces longer-term federal and provincial forecasts, while labour market analyses (LMA) identify and continually adjust to current regional and short-term trends.

Workforce projections are commonly used at the federal level to provide long-term estimates (Van Adams, Middleton & Ziderman, 1992). However, estimates are only as good as the "plausible assumptions" on which they are based. Because there are many unforeseeable factors that may affect the economic growth of a country (and the labour needs this growth produces) the quality of forecasts tend to decrease as the length of the forecasting period increases (Campbell, 1997a). Projections in terms of sectoral change, on the other hand, are often reasonably accurate. The difficulty here appears to be translating the accompanying skills change into a profile that can be used to inform decisions and policies regarding looming training and education requirements (Psacharopoulos, 1991).

Labour Market analysis uses "signals" to forecast future labour and education requirements. These signals are available via newspaper job listings or from the provincial and public employment and social insurance services that collect information about job openings, placements and unemployment rates. These resources provide a wealth of data which, if analyzed, may offer insight into current shortages or surpluses of workers. This approach can be limited by the fact that the data gathered from newspaper listings and public employment services are not generally available in a form that can easily be analyzed, and are unlikely to be complete since many jobs are not posted and many unemployed do not register with public services. Other types of signals include employer and household surveys, enrolment data and tracer studies, all of which tend to be restricted by region and population.

In total, 38 studies were included in this review, 28 of those used empirical strategies to demonstrate the accuracy of a particular forecasting model. Our analysis and evaluation of these studies resulted in a two major conclusions. First, the quality of the literature devoted to labour market forecasting is inconsistent, meaning readers and reviewers of such material need to be proficient in econometric modelling and research design if they are to fully assess the value and or the flaws within the conclusions drawn by the authors. Second, forecasting research is very much source, location, and time specific. Consequently, it is not clear if the models will perform as well in other forecasting horizons. Our study suggests that there is no single forecasting model that can accurately forecast labour market needs in all situations. While some of the proposed models show an impressive level of accuracy in forecasting

within a particular market, without controlled replicability the consistency of the forecast accuracy remains uncertain.

Ultimately, it appears that some forecasting models have the ability to estimate labour needs in very specific circumstances. What remains unknown is whether it is possible to develop a single model that will accurately forecast in a range of situations, under various conditions.

1.0 BACKGROUND

1.1 Introduction

The purpose of this review is to systematically and transparently gather, analyze and synthesize both empirical research and peer-reviewed analytical articles devoted to discussing and determining whether it is possible to accurately predict labour market needs. A variety of academic databases, websites and key research journals were searched, resulting in an initial capture of 3,413 articles. The application of specific inclusion and quality assessment criteria reduced this number to 36. Eight of these articles are peer-reviewed analyses of labour market forecasting, while the remaining 28 are econometric investigations that aim to improve the accuracy of labour market forecasting. Both sets of articles have been analyzed to produce a comprehensive synthesis of information on how to maximize the accuracy and utility of labour market forecasts.

This review is divided into five sections. In the first section, we frame the issue, placing it within the Canadian context and outlining the different approaches to forecasting. In the second section, we discuss in detail the methods employed in this systematic review. The third section includes précis of the 36 articles included for final analysis. The fourth section is devoted to discussions of the quality of the research. A rubric of quality is included to ensure the transparency of the scoring process. In the fifth section, we synthesize the literature, and discuss the results and further considerations of our review.

1.2 Why Attempt to Forecast Labour Market Needs?

Attempts to forecast the labour market began after the end of WW II, fuelled by the economic growth occurring at the time and the belief that, through planning, governments could ensure that the supply of available workers were trained or educated to fit the needs of the production system (National Observatory, 1999). Since its inception, forecasting has been the subject of much debate among economists.

The poor track record of forecasts and the complex task of "tuning" the labour market (Heijke, 1996) have led to the view that labour market planning is not necessary nor is it useful (Borghans & Willems, 1998). Those subscribing to this view believe that the market will correct itself as individuals and employers respond to labour market signals and decide what sort of skills merit the investment of training or increased wages (Haskel & Holt, 1999; Neugart & Schomann, 2002; Psacharopoulos, 1991). However, there are strong arguments to support the practice of looking ahead, despite the difficulty in predicting all the variables that may affect the future labour market.

As Venari (1999) points out, one consideration to be aware of is how "shortage" is defined. A shortage of teachers, for example, can simply mean that there are not enough trained teachers to fill the number of spaces necessary to provide adequate education. Or it can mean that there are not enough teachers or potential teachers willing to work at the current wages provided for that work. These two definitions present different issues. Some argue that the market can adjust to "shortages" itself by raising wages, but only the second type of shortage can be addressed through wages.

Indeed, the market does adjust to some degree, and employers do raise wages to attract more workers. Higher wages, in turn, create an incentive for individuals to invest in training. This process can lead to what Heijke (1996) terms the "cobweb cycle," where students base their educational decisions on the market at the time they enter a course, rather than the market anticipated at their time of graduation. For example, if wages are high in a certain field due to a shortage of employees, many students may respond by taking the courses necessary to enter that field. After a number of cohorts have done this, there is a surplus of employees and wages go down, causing new students to stop entering that field. At the end of the cycle, a new shortage in labour occurs, and wages again go up to attract more students. Accurate information about the labour market that is frequently updated and widely disseminated can help to avoid the pitfalls of this cycle (Heijke, 1996).

Another reason that it may be unwise to leave the market to correct itself is the fact that employers also use wages for other purposes, such as retaining workers with seniority, reducing turnover or increasing worker productivity (Neugart & Schomann, 2002). Also, many employers may be unable to raise wages due to a "fixed compensation structure" in their organization (Veneri, 1999, p. 17). Raising wages can also reduce the international competitiveness of firms.

Some argue that employers should be responsible for providing education and training to meet their needs, rather than benefitting from government intervention. But for employers, investing in training when there is a tight labour market is risky, as other firms may "poach" newly skilled workers (Haskel & Holt, 1999). Instead, employers often compromise by reducing the skill level required to fill certain positions, thus lowering quality and creating disincentives for individuals to invest in their education. This situation can lead to a "low skills/low wage equilibrium" in the market, which is difficult to override (Haskel & Holt, 1999; Neugart & Schomann, 2002).

Another complication is that both firms and individuals simply may not recognize the signals of changes in the labour market; slow response time may delay market adjustment. Workers may choose to remain underemployed or even unemployed and "wait out" a downturn that may be only temporary. Even if workers do respond to a perceived shortage, training institutions may not be able to accommodate the demand without the benefits of foresight and policy support (Venari, 1999). Adjustment is costly and takes time (Smith, 2002).

Forecasts and policy interventions are unlikely to completely eliminate the cycle of skills supply and demand since the labour market is constantly evolving and adjusting. They do, however, help to eliminate the "firefighting" approach to the labour market by enabling strategic planning for upcoming shortages and surpluses, which can mitigate the costs of slow adjustment (Neugart & Schomann, 2002, p. 2). For most individuals, the labour market is the single most important market in their lives, and any understanding that allows for improved efficiency can create significant social benefits (Smith, 2002).

Forecasting provides policy-makers, employers, employees and students with the information necessary to make choices that will optimize the contribution of education to their economic growth and the smooth functioning of the labour market (Heijke, 1996; National Observatory, 1999; Neugart & Schomann, 2002). Better-educated people are at a lower risk of unemployment and generally obtain better jobs at higher salaries. In addition to

promoting economic growth, investment in education is a social good, reducing crime rates and increasing social cohesion (Neugart & Schomann, 2002). Planning ahead also allows policy-makers to take into account issues of equity and poverty, and work to adjust the labour market to make improvements in these areas (Psacharopoulos, 1999). Providing and disseminating "labour market information is a public good in that many users can share the same information and the benefits of its production are equally available to non-payers" (Smith, 2002, p. 68).

This information is especially important when there are long "lead times," such as those for training in specialized and necessary skill areas such as medicine, technology or teaching. Shortages in skilled labour in these areas could conceivably contribute to longer-term social and economic problems, and forecasting helps reduce this risk (Neugart & Schomann, 2002). Understanding how the labour market should be directed in order to meet social and economic objectives helps governments monitor and influence both the labour and education markets in a manner consistent with these objectives (Borghans & Willems, 1998). Such foresight improves the efficiency and cost-effectiveness of the labour and education markets.

Advance knowledge of where to expect skills shortages and surpluses allows governments, individuals and employers to invest in education that will maximize the return on their investment and helps prevent the loss involved in training people in skills that are no longer in demand (Burns & Shanahan, 2000; National Observatory, 1999; Neugart & Schomann, 2002). Making this information available to career counsellors, students, the unemployed, employers and those attempting to change careers can facilitate informed educational decision-making. Ideally, this information also aids group decision-making since all interested parties are able to make training and employment decisions based on a common set of assumptions about where the labour market is heading (Volterra Consulting, 2005).

However, while labour market forecasts can contribute to informed decisions at the individual and policy level, they are only as good as the information upon which they are based. Projections that are based on faulty assumptions about the future will be inaccurate and will lead to mistaken predictions of demand in certain sectors. Different types of predictions are suited to different purposes and audiences, and require different inputs of information. Perhaps most relevant to the issue of training and education, inaccuracies can arise from mistaken assumptions about what skills are needed for different occupations. These problems and other issues involved in accurate prediction of labour market needs will be discussed in the following sections.

1.3 The Canadian Context

Canada's well-established history of labour market forecasting began in the 1960s. Canada currently uses a forecasting model called the Canadian Occupational Projection System (COPS) (Archambault, 1999; Willems, 1996). COPS was developed by the Applied Research Branch of Human Resources and Social Development Canada (HRSDC), which is the main supplier of occupational forecasts in Canada and is supported by a system of information provided by federal, provincial and territorial researchers and analysts. The methodology has evolved over the years to ensure the continuous improvement of the accuracy of the information produced by COPS (Archambault, 1999). Originally, COPS was a strictly

"demand side" model, but since the 1990s, the model integrates supply and demand. It provides forecasts for 139 occupations and five broad skill categories, and is updated every five years (Neugart & Schomann, 2002; Smith, 2002). COPS is supplemented by short-term regional and sectorial forecasts (Neugart & Schomann, 2002). Both a peer-reviewed analyses of COPS and an empirical evaluation of the strengths and weaknesses of COPS and other short-term regional Canadian models are included in Section 3 of this review (see Archambault, 1999; Fauvel, Paquet and Zimmerman, 1999; Smith, 2002).

The results of COPS are available for widespread use by Canadians through HRSDC's publication *Job Futures*, which is posted on the internet (http://jobfutures.ca/en/home.shtml). Career counsellors, individuals and institutions use this information to inform educational planning, and to develop curricula and programs. Business and industry use it to develop occupational standards, and assess training and human resource needs. The federal government uses it to create training programs in appropriate areas (Neugart & Schomann, 2002). Many provincial governments use the COPS data system to create comparable provincial-level forecasts that take into account the specific demands of industry in their province (Smith, 2002).

1.4 Approaches to Forecasting

There are two major approaches to occupational forecasting: workforce projection and labour market analysis (signalling). Workforce projection produces longer-term federal and provincial forecasts, and should be supplemented by labour market analysis approaches (LMA), which identify and continually adjust to current regional and short-term trends. The federal COPS system begins with a series of macroeconomic forecasts of GDP, and provincial forecasts based on COPS often include LMA specific to the future of the labour market in that region. This section will briefly explain the methods used in these approaches, and the strengths and weaknesses of each. Most countries and provinces have developed unique approaches to maximizing the accuracy of the information that goes into, and is produced by, their models.

1.4.1 Workforce Projection and Forecasting Approaches

Workforce Planning

The COPS system is an example of the workforce planning approach, as are the Bureau of Labour Statistics projections used in the USA. This approach provides longer-term forecasts that relate the output of training institutions to a country's planned economic growth by a target year, based on "plausible assumptions" (Campbell, 1997a, p. 281). The employment levels necessary to achieve that growth by the target year in each economic sector are forecast, and the distribution of occupations in each sector is estimated. For each occupation, these forecasts are converted into projections of labour requirements based on the imbalance between expected occupational growth and replacement[1] needs (demand) and expected available new labour in the form of graduates, immigrants and labour force re-entrants (supply). This provides an estimate of labour requirements. The estimate can then

[1] Replacement refers to the need for new workers due to retirement, promotion, job turnover, death, career change, etc.

be used to determine training requirements based on the skills and education needed for each occupation. The estimated training requirements are then compared to the expected output of training institutions, enabling planners to identify potential shortages or surpluses (Archambault, 1999; Campbell 1997a; Van Adams, Middleton & Ziderman, 1992; Willems, 1996).

This general method is commonly used at the federal level to provide long-term estimates because it is relatively practical and straightforward, and has reasonable data requirements (Van Adams, Middleton & Ziderman, 1992). However, there are many issues that can jeopardize the reliability of this method. To begin, the estimates are only as good as the "plausible assumptions" on which they are based. Unfortunately, there are many unforeseeable factors that may affect the economic growth of a country and the labour needs this growth produces (Campbell, 1997a). One commonly cited example of such factors is the rise of the PC and internet in the 1980s and 90s, which created a skills demand that could not have been predicted. Also, this method assumes constant wage relative rates and a fixed relationship between the number of workers needed and the quantity of goods produced, when in reality there are many ways in which labour and technology can be configured to increase or decrease productivity in certain sectors (Smith, 2002; Van Adams, Middleton & Ziderman, 1992). The labour market, as noted above, will find many ways to adjust to a skills shortage, which may eliminate the need for certain jobs. Because of these variables, the quality of the forecasts decreases with the length of the forecasting period.

Projections in terms of sectoral change, on the other hand, are often reasonably accurate. The difficulties arise in translating the occupational and labour requirements into educational and skill profiles that can inform policy on training and education (Psacharopoulos, 1991). In Canada, Statistics Canada classifies occupations using the Standard Occupational Classification (SOC) coding system, which begins with 10 broad occupational categories conceptually based on "homogeneity with respect to skill type" (Statistics Canada, 2006). The SOCs are converted into National Occupational Classifications (NOCs) and provided to HRSDC for their forecasts (Archambault, 1999; Statistics Canada, 2006). NOCs contain information on employment requirements for an occupation, they indicate typical occupational progressions, and they are subdivided by skill types and skill levels (Statistics Canada, 2006). The emphasis on skills may be useful, as a major challenge in determining training needs is the fact that most skill sets can lead to a number of different occupations; in practice, "people arrive in occupations by diverse paths" (Van Adams, Middleton & Ziderman, 1992, p. 264). Planning forecasts are often criticized for ignoring the substitutability of various types of educated labour.

Another limitation of these types of forecasts is the assumption that the skill set required for a given occupation at the time of the forecast will be the same as that required in the target year. Bailey (1991), whose analysis of occupational forecasts is discussed in Section 3, points out that many positions that once required only high-school education now require college or university education. In forecasting change, it is risky to assume that education and skill requirements will remain the same. Recently, there has been a rise in jobs that involve multi-tasking, which blurs the line between skills and occupations. Changes in skill requirements do not necessarily equal changes in occupations (Burns & Shanahan, 2000; Haskel & Holt, 1999). These serious limitations have led many of the authors cited here to suggest that workforce planning forecasts with shorter time spans are more appropriate, and that these

forecasts need to be updated regularly (Campbell, 1997a; Heijke, 1996; Psacharopoulos, 1991; Willems, 1996).

Econometric Models

We review a number of econometric models in Section 3 of this report. Econometric models provide an alternative method of performing the first step in workforce planning (i.e., forecasting economic growth by a target year as a basis for determining the employment levels needed in each economic sector to reach that growth). Indeed, econometric models are often used to complete this first step, and the COPS model incorporates input from many models to estimate overall GDP growth (Smith, 2002). Econometric models are complex mathematical representations of the interrelationships between economic variables and sectors (Burns & Shanahan, 2000; Campbell, 1997a); however, they still must rely to some degree on economists' theories about causal relationships in the economy (Haskel & Holt, 1999).

In theory, there are an infinite number of relationships that can be modelled, but, in practice, the infinite data necessary to model these relationships is not available (Haskel & Holt, 1999). Also, as Burns and Shanahan (2000) point out, "there is a trade off between the number of relationships that can be modelled (level of disaggregation) and the complexity of behaviour that can be modelled" (p. 14), so simple mathematical characterizations of decision-making may need to be used in conjunction with models that attempt to capture the relationship between a large number of variables. As a result, most economists simplify the models, choosing their modelling methodologies based on the questions asked and answers sought, and/or splitting the modelling process into parts (Burns & Shanahan, 2000; Haskel & Holt, 1999). One of the strengths of these models is their dynamic nature; databases can be updated with new information that directly affects employment projections (Campbell, 1997a). Theoretically, these models can also be used to map out different scenarios in the economy (Haskel & Holt, 1999).

Burns and Shanahan (2000) outline the desirable properties of sound econometric models. First, these models should be consistent, so that predictions from one part of the model will not contradict those from another. Second, the supply and demand relationships should be based on sound economic theory. Third, in order for the models to accurately predict real-world outcomes, the "parameters that reflect the sensitivity of quantity variables to changes in prices and income" (Burns and Shanahan, 2000, p. 16) must be estimated from market data using statistical methodology. Burns and Shanahan also suggest that the models be replicable and provide explanatory documentation that makes them accessible and comprehensible to users. Lack of comprehensibility is one of the main drawbacks of econometric models (Campbell, 1997a).

There are other limitations to these models. Some of these limitations are similar to those associated with workforce planning: assumptions are part of the input into the models; "shocks" to the economy are usually unforeseeable; there are variables that are outside the control of governments; and it is difficult to accurately model the decision-making behaviour of the individuals in the labour force. Also, econometric models require large amounts of data and the model databases are costly to maintain. The data tend to be national, so that the results may not be generalizable regionally. Like any other predictive instrument,

econometric models are subject to error, and, unfortunately, modellers cannot provide information about confidence intervals for the outcomes predicted. Nonetheless, these models generally can provide higher-quality forecasts than other methods, as long as they begin with quality data (Burns & Shanahan, 2000; Campbell, 1997a; Haskel & Holt, 1999).

1.4.2 Labour Market Analysis Approaches

Multiple measures from different perspectives on the labour market should be included as a supplement to forecasting techniques to increase their accuracy and regional and sectoral applicability (Vanari, 1999). Labour market analysis refers to a variety of measures of education and training requirements, which can be used to translate occupational classifications into actual skills requirements. Concern for issues of poverty and equity are also prominent in labour market analysis (Psacharopoulous, 1991). Taking these measures into account can strengthen the relationship between forecasters and all beneficiaries of forecasts, and improve the sharing of information and thus the forecasts themselves (Campbell, 1997a; Van Adams, Middleton & Ziderman, 1992).

Public Employment Services/Job Advertisements

One type of "signal" from the labour market is readily available from newspaper job listings or from the provincial and public employment and social insurance services that collect information about job openings, placements and unemployment rates. These resources provide a wealth of data that, if analyzed, offer great insight into current shortages or surpluses of workers. High levels of unemployment may indicate a shortage of available work or a need for subsidized retraining in skills that are in demand. Rises in the salaries offered in certain sectors or vacancies that remain unfilled for a long period of time may be signals of a shortage of skills in that area. If analyzed over time, this information could also begin to reveal trends in the labour market (Campbell, 1997b; Van Adams, Middleton & Ziderman, 1992). Job advertisements also contain concrete information about the skills and qualifications that employers are looking for in certain positions, which can contribute to occupational profiles.

This approach is limited by the fact that the data gathered from newspaper listings and public employment services are not generally available in a form that can easily be analyzed, and are unlikely to be complete since many jobs are not posted and many unemployed do not register with public services (such services are usually located in urban areas and may not offer insight into the rural labour market). In addition, these data cannot be used to predict the future of the labour market (Campbell, 1997b).

Key Informant Interviews

Key informants are individuals who are in a position to know a great deal about certain sectors, educational areas or economic trends. These people can offer information about the types of skills in demand in upcoming years, the scarcity or abundance of certain skills, anticipated technological advances, and hiring practices in certain occupations.

This technique is inexpensive and, provided the informants are chosen carefully, can be very useful for wide variety of forecasting purposes, both regional and national. However, the

data collected is qualitative, and informants may have biased opinions about the labour market (Campbell, 1997b).

Employer Surveys/Household Surveys

Employers are among the potential beneficiaries of labour market forecasts. Since they make hiring decisions, they can also provide input into labour market forecasts regarding upcoming skills needs, the skills and education profile needed for certain jobs, and the types of businesses that are growing. Household surveys can supplement employer surveys by providing information on those who are not in the organized labour market (the self-employed, unemployed, stay-at-home parents, students, etc.) but who may be providing labour through contract work or planning to re-enter the workforce in the near future (Psacharopoulos, 1991).

This technique goes "right to the source" (Campbell, 1997a, p. 287) of the supply and demand sides of the labour market, supplementing information gleaned from other sources and further informing assumptions about the labour market. Surveys and face-to-face interviews (depending on the type of data needed) with employers and members of households can provide in-depth information, especially at the regional level and in small communities (Campbell, 1997b). However, qualitative information, as noted above, is subject to bias and may not be applicable to other regions or sectors. Also, surveys sent to households and employers may not have a high a rate of return (Campbell, 1997b).

Enrolment Data and Tracer Studies

Identifying courses that have high or low applicant-to-admission ratios provides indications of areas of high or low demand for specific types of training. This technique generates information that can guide educational investment decisions on the part of individuals and businesses; it also may be used to predict the expected number of trained workers in certain areas in future years. Many training institutions also conduct tracer studies, which follow up on graduates after they enter the workforce to discover where they have found employment. Tracer studies can indicate the likelihood that graduates will actually work in the field for which they are trained, the sectors that are hiring graduates with certain skill sets, the length of time it takes to find a job in a specific field, the wages paid in different occupations and the mobility of the workforce (i.e., the distance graduates had to travel in order to find work). Tracer studies enable planners to create a "map" of the different routes into an occupation (Van Adams, Middleton & Ziderman, 1992; Campbell, 1997b; Psacharopoulos, 1991).

Enrolment data may be misleading, as only a percentage of enrollees will stay in one program, graduate and immediately enter the workforce in a directly related field. Tracer studies also have drawbacks: they may have other objectives that are not directly related to the labour market, and they may suffer from low response rates if they are too lengthy or if graduates have relocated. Nonetheless, these data are often readily available from training institutions in a form that can be easily analyzed (Campbell, 1997b; Psacharopoulos, 1991; Van Adams, Middleton & Ziderman, 1992).

1.5 The Body of Research Reviewed

The goal of accurately predicting labour market needs is a popular one. The literature in this area spans a wide range, from general theoretical discussions of accurate predictions to specific econometric models purported to solve prediction challenges in one small area. Some of the literature focusses on the training and education implications and the social benefits of labour market planning, while some focusses solely on maximizing productivity and economic growth.

There is general consensus in the literature that complete accuracy in forecasting is not possible because the task requires too many assumptions. Our review of the literature focusses specifically on the body of work that directly discusses the issues surrounding accurate prediction of labour market needs or empirically tests a model to determine its utility in accurate prediction.

1.6 Review Question and Approach

This review evolved from an earlier, more limited, search designed to determine the feasibility of conducting an in-depth review on this topic. The tool used for this search is known as a Question Scan.[2] Question Scans survey the literature devoted to a given topic using a sample of relevant databases and fugitive literature websites. The purpose of such a survey is to determine the feasibility of carrying out a much larger, more comprehensive, systematic review of evidence on a given question.

Given the results of the Question Scan, the British Columbia Ministry of Advanced Education decided to proceed with a systematic review of the question, "Is it possible to accurately predict labour market needs?" Through discussions with ministerial staff, a number of elucidations were made regarding the scope of the question. The aim of this review is to address, where feasible, the possibility of accurately predicting growth in employment in different sectors and regions in order to ensure that government policy can be implemented with foresight, thus preventing skill shortages or surpluses. A large body of literature was also available regarding the means to accurately predict labour market supply and demand (rather than the larger question of whether or not it is possible to do so). This literature was determined to be outside the scope and means of this review.

A variety of databases across disciplines were searched. The disciplines include business, education and economics, and are identified in greater detail in the methodology section. A list of professional organizations with reputations for conducting research in the field was also searched.

Inclusion/exclusion criteria were developed by the review team and approved by the client. The retrieval of documents identified for secondary review was conducted using internet sources and a hand search at the UBC library. Final inclusion criteria were developed by the review team and approved by ministerial staff.

[2] CCL would like to acknowledge EPPI as the developer of the original version of a Question Scan.

Because the majority of the studies included for review used econometric methods, an econometrician was added to the team to carry out the analyses of the studies and their quality. A rubric was developed (and has been included[3] to ensure transparency) for scoring the quality of the studies.

Graphs and tables are included where possible to enhance the summative findings, and add clarity and ease of reference through visual representations.

1.6.1 Being Systematic

Essential components of systematic reviews are documentation and transparency. Protocols and procedures are followed to ensure an orderly and organized approach to the review process.

CCL's approach to conducting reviews of evidence follows well-established review protocols: thorough and transparent procedures are used to collect and analyze evidence-based research in education. CCL's research and review teams use a variety of documentation templates when conducting Question Scans and systematic reviews to ensure consistency, transparency and accountability.

The use of the above-mentioned templates allows for the replication of our reviews. Each review conducted by CCL includes the use of a search diary by the information retrieval specialist who conducts the searches. In the same fashion, all inclusion/exclusion criteria and decisions are documented, as are reasons for any exclusions. CCL has also developed a document control sheet to track and monitor the movement of studies from their initial capture through to their final inclusion and analysis. A flow chart illustrating the movement of studies through the review is also provided.

1.6.2 Stages of the Systematic Review

This section provides a short description of each stage of the systematic review process. The results of each stage and their subsequent application are described in greater detail in the methodology section which follows.

Identification of the research question(s) and conceptual framing

Consultation with the client to identify key concepts and emphases of the research is the first step of the review process. At this stage, terms used to formulate the research question are clarified and defined, and relevant databases and websites are identified with the assistance of experts and practitioners in the field.

Development of search strategies

CCL created a search strategy development process, which includes the creation of a comprehensive list of both free terms and controlled vocabulary. This process ensures a broad, yet focussed, capture of literature devoted to a particular question. Free terms are generally provided by experts or practitioners in the field, and are used to identify matching

[3] See Section 4.2

and/or complementing controlled vocabulary using database thesauruses. A facet analysis (which involves grouping like terms according to category) is performed for all terms. Search terms are then linked using truncation and Boolean terms, resulting in a single search strategy for each category of terms (i.e., population/location, issue and method). All category search strings are then combined to create the final search strategy. Search strategies will differ slightly according to the capacity of the database.

Searching

CCL's search procedures include the use of a search diary. The diary is used to document all searches, their results and any decisions to adapt the search. Search results are imported to a database and sorted for duplicate references.

Searching for fugitive literature (research not published in journals) requires the collaboration of experts and practitioners in the field. Key websites are gathered and reviewed for relevant publications, references, and links to other pertinent sites.

These websites are then used to locate similar websites using the advanced search function of Google, "find pages similar to the page." The reference lists of included articles are also reviewed for relevant studies.

Initial review of results

Once all the search results have been imported to the reference management database and sorted for duplicates, inclusion and exclusion criteria are developed to remove irrelevant material from the review process. The criteria are developed collaboratively by the review team and sent to the client for editing and approval.

The initial inclusion/exclusion criteria are applied using titles and abstracts only. Reviewers are instructed to include studies in cases where it is difficult to decide whether they meet the criteria so as not to eliminate relevant studies in error. Exclusion decisions are recorded in the inclusion/exclusion spreadsheet.

Document retrieval

Once the initial inclusion/exclusion stage is complete, the full texts of all included articles are retrieved for the second stage of review. The majority of the studies are obtained via the internet. The reviewer saves an electronic copy of each retrieved study. Studies that are not available via the internet are located through university libraries whenever possible. The information retrieval specialist tracks the retrieval of the studies using the document control sheet. Each study is tagged with an identification of its origin, which may be a database link, the internet (typically Google scholar), the library or inter-library loan. *Secondary review*

Secondary inclusion criteria are typically more detailed than the initial criteria because the reviewer has acquired a general sense of the body of literature owing to the initial review of the titles and abstracts. It is at this stage that more studies are excluded due to inappropriate

research designs or a lack of focus on the specific question at hand. Each included article must meet specific inclusion criteria.

Summaries and quality analyses

Each study is summarized by the reviewer. In the case of this review, an econometrics expert was brought in to review and analyze the quality of the included studies. A rubric of quality was developed, applied to each study and included in the report.

Synthesized findings

Studies are synthesized using a variety of techniques and methods appropriate to the question and the resulting capture of studies. Using the coding sheet, studies are grouped and synthesized according to focus, results and study quality. General statements of findings are made accordingly.

2.0 METHODS

2.1 Search Strategy

The search strategy for this review was developed by the research team in consultation with experts in the field, including British Columbia Ministry of Advanced Education staff. Prior to the development of a search strategy, the research team reviewed key background articles located by the initial Question Scan. These articles helped to inform our understanding of the issues and terminology relating to the accurate forecasting of labour market needs. Our initial search was structured broadly in order to maximize the capture of articles. The search terms included, but were not limited to, the following:

- Issue 1: economic development, economic growth, employment, employment level, human capital, human resource management, industry profiles, labour demand, labour market, labour needs, labour supply, manpower policy, skilled workers, skilled occupations

- Issue 2: business forecasting, economic forecasting, economic research, employment forecasting, employment projections, estimation theory, forecasting techniques, needs assessments, planning, prediction models, trend analysis

- Method: models, data analysis, data collection, systems approach, econometrics, measurement technique, environmental scanning

Using these category terms, controlled vocabulary were collected and final search strategies developed for each of the following nine databases: ERIC, EconLit, Academic Search Premier, Business Source Premier, Theses Canada Portal, Education Full Text, CBCA – Education, CBCA – Business, and CBCA Dissertation Abstracts. A date limiter of 1985 to the present was imposed on all searches. In total, 3,347 articles were captured.

In addition to the database searches, we performed a hand search. The bibliography of our original Question Scan yielded 11 articles. The bibliographies of all the retrieved articles were also reviewed for additional relevant articles. In addition, a fugitive literature search was conducted in key education research websites in order to capture relevant research not published in academic journals. The following websites were searched:

US Dept. of Labor Bureau of Labour Statistics: http://www.bls.gov

Ministry of Labour and Citizens' Services BC: http://www.gov.bc.ca/lcs/cont/

Gov. of Canada Human Resources and Social Development: http://www.hrsdc.gc.ca

Statistics Canada: http://www.statcan.ca

Department of Labour New Zealand: http://www.dol.govt.nz

Gov. of Alberta Human Resources and Employment: http://www.hre.gov.ab.ca

Canadian Policy Research Network: http://www.cprn.org/

Warwick Institute for Employment Research: http://www2.warwick.ac.uk/fac/soc/ier/

European Training Foundation: http://www.etf.europa.eu/

Conference Board of Canada: http://www.conferenceboard.ca/

Construction Sector Council: http://www.csc-ca.org

Canadian Steel Trade and Employment Congress: http://www.cstec.ca/

Canadian Labour and Business Centre: http://clbc.bc.ca

The Alliance of Sector Councils: http://www.councils.org/

Lastly, an internet search was conducted in order to identify any additional relevant fugitive literature, using the following terms: forecasting models, labour markets, skills shortages, and education, training or vocation.

All citations were uploaded to a single citation management account. The final total of studies initially captured was 3,413. After sorting for duplicates, 3,242 remained and were included in the initial inclusion/exclusion process.

2.2 Inclusion and Exclusion Criteria

As noted earlier, reviews of evidence are collaborative endeavours. The development and finalization of inclusion and exclusion criteria were completed with the assistance of ministerial staff. At the end of the first inclusion process, we had included 426 articles, a number considered unmanageable given the time/resource constraints of this project. We reset our date limiter to include only articles published after 1990, which reduced the number of included articles to 323. This was deemed a large enough sample to draw valid conclusions.

The final criteria for exclusion of articles fell within five categories:

1. *Temporal*: Studies published prior to 1990 were not included.
2. *Population*: Studies conducted in non-OECD countries were not included.
3. *Issue*: Documents that were not about forecasting labour market need, supply or demand, and studies that included identified keywords but were not relevant to the question were not included.
4. *Intervention*: Studies that did not assess a forecasting model were not included.
5. *Methodology*: Books, book reviews, magazine articles, editorials, and studies that were not peer-reviewed were not included.

Attempts were made to retrieve all 323 studies; however, time and resource constraints did not permit the retrieval of articles that were not accessible online or at local university libraries. These constraints resulted in the elimination of an additional 50 studies.

2.3 Mapping and Summarizing

Full reports of the remaining 273 articles were read and secondary inclusion criteria applied. This process enabled the review team to identify those studies that meet a minimum empirical standard of methodology and data analysis for evidence-based research. Because this review seeks to determine whether it is possible to accurately forecast labour market needs, it was also deemed important to include articles from experts in the field that directly address this issue, although they may not use strictly empirical methods.

In order to proceed to the final analysis and coding stages of the review, a study was required to meet one of the following criteria:

- It is an analytical/critical or descriptive paper that directly addresses the topic of accurate forecasting of labour market needs and is published in a peer-reviewed journal.
- It empirically tests a forecasting model to determine its accuracy and utility in predicting labour market needs, supply or demand, and it includes results of the forecasting that are relevant to the question.

Articles that did not meet the inclusion criteria but provided relevant background on labour market planning were also earmarked at this time. These included relevant reviews of literature, criticisms of forecasting models, background on forecasting methods and their limitations, meta-analyses, and reviews that dealt with the importance of predicting labour market needs. These studies were used to inform the direction of the report.

Sixty-seven studies met the secondary inclusion criteria: 36 of these studies were included for further analysis, while 31 were identified as "background material" (i.e., studies that provide contextual, historical or theoretical information but offer little in the way of analysis or data).

Studies were separated according to the type of research: peer-reviewed descriptive/analytical/critical (DAC) or empirical. DAC studies were summarized, and any critiques or endorsements of forecasting attempts or models of forecasting skills shortages were highlighted. Précis were written for the empirical research. Most of the empirical research consisted of econometric studies that proposed a new forecasting model or tested an existing one. A summary table identifying the important characteristics of each study was created.

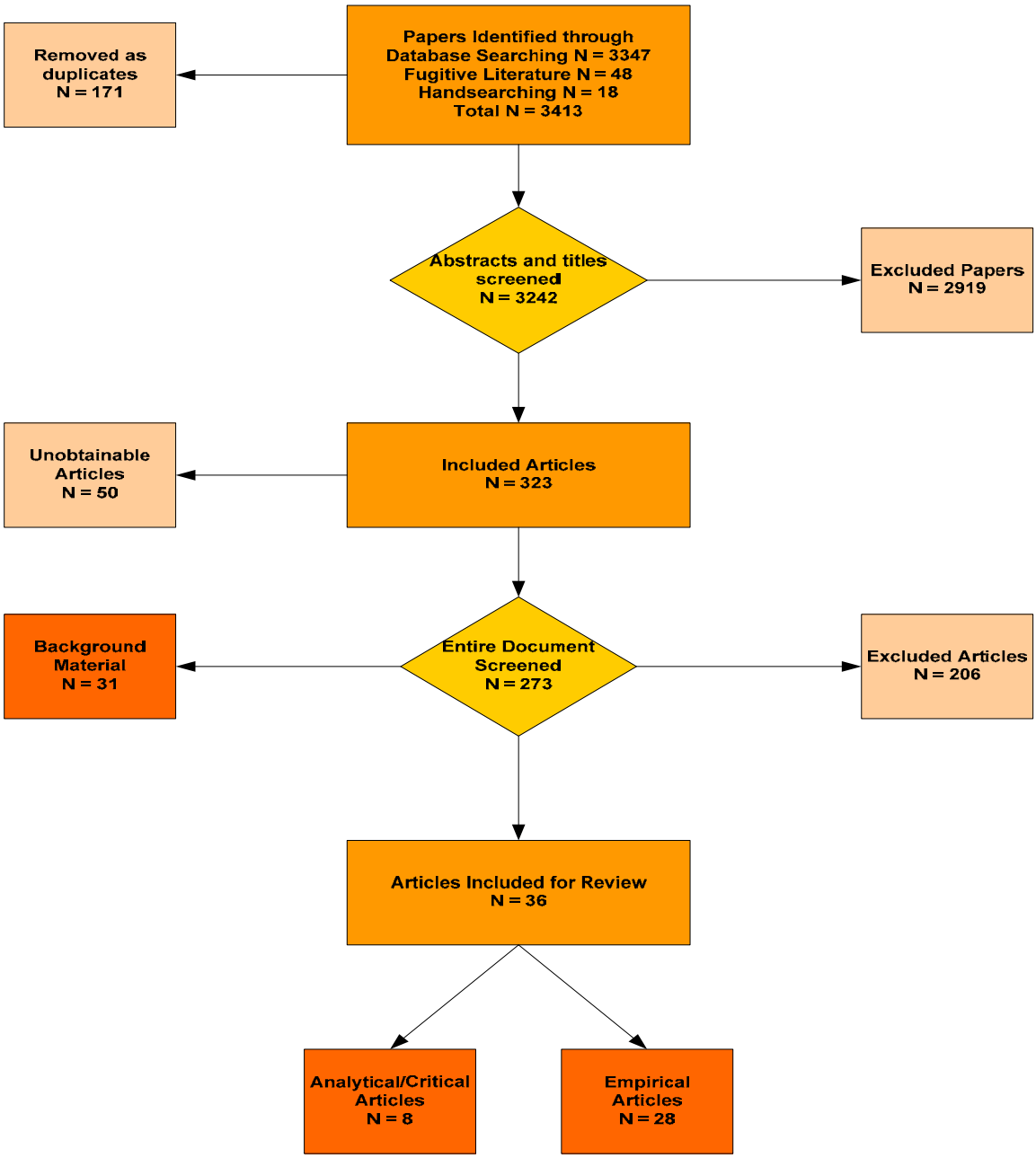
2.4 Review and Analysis

Due to the large number of articles included for analysis and the complexity of the econometric designs used in the empirical studies, CCL brought in an econometric expert to assist in the analysis of the research. A rubric was developed for evaluating the quality of the research designs and applications. Each study was summarized and analyzed by the econometrician and awarded a score ranging from 11 to 33 based on the rubric criteria. The results of this process are described in greater detail in Section 4.2.

3.0 Mapping the Results—Identification and Description

3.1 Flow of Literature through the Review

In Sections 1 and 2, we outlined the systematic approach adopted to search and identify relevant articles. The flow chart below is designed to make the progress of literature through the review process transparent.



3.2 Peer-Reviewed Qualitative Studies

3.2.1 Introduction

The articles included below are descriptive, analytical or critical articles, from peer-reviewed journals, books or research institute publications that address the issue of whether or not it is possible to accurately forecast labour market needs. The majority of these articles deal with existing workforce forecasting approaches such as the COPS system or the United State's BLS projections. The frequent criticisms of these approaches are addressed. Various sources of error are identified, and solutions are proposed. The sources of error include confusion regarding issues of labour market theory, vague definitions, faulty assumptions and data limitations. The authors examine the utility of forecasting and the means to improve its efficacy through interpretation and dissemination of results.

3.2.2 Précis of qualitative studies

Heijke (1996)

We chose to begin with Heijke (1996) because he provides a broad background on policy approaches and labour market theories, which supplies a useful context for this report. Heijke advocates increased transparency and relevance when providing labour market information to those with an interest in educational investment, so that the supply side of labour can respond to market changes more efficiently. He repeats the frequently mentioned difficulties related to forecasting. First, he points out the difficulty in measuring the "productive yield" of training and experience, since productivity is "the result of an interaction between the occupation and the person filling the occupation" (p. 2). Heijke criticizes the assumption, inherent in most forecasting approaches, that "pay is equal to the marginal productivity of labour," because wage data is usually incomplete (it does not differentiate by age, gender, experience, education, occupation or sector). He also points out two further sources of confusion: there is no consistent understanding of the functioning of the labour market or the specifics of the relationship between education and labour markets, and even if there was, there are always circumstances outside the labour market that can have unforeseen impact.

Heijke outlines various policy approaches to coordinating the labour market and education. One of these approaches, educational investment based on workforce forecasting, often fails because of the faulty assumption of an "exclusive relationship between job requirements and training" (p. 4). In response to this failure, the policy approach of broadening initial training courses so that they can lead to a number of occupations has been promoted. This approach ensures a workforce that can adapt quickly: Specific skills can be picked up through shorter courses or on-the-job training. The drawback is that, with a broadly based education, workers do not have much time to develop the specialized skills of their occupation and productivity may suffer. The third policy approach, moving away from direct intervention in education, has been to improve the "adaptability of the supply side of the economy" (p. 5) through the provision of information that enables educational investors to respond to labour market fluctuations.

Heijke presents research indicating that students do pay attention to the labour market when making educational investment decisions but that their understanding of the market is not always accurate. He argues that this inaccuracy is caused by the incomplete supply of labour market information and the fact that the available information is often not relevant to students' educational choices. Improving this information may help eliminate the "cobweb cycle," described in the introduction of this report. Essentially, Heijke advocates an optimal combination of all three policy approaches: broaden and shorten initial training, conduct forecasts and disseminate all relevant information available to those who are making the investments.

However, the choice of "relevant" information depends on the theory one adopts regarding the functional link between the labour market and educational investment. Human capital theory assumes that people choose occupations that have the highest benefits and lowest costs. Thus, information about wages for certain educational backgrounds would be highly relevant to decision-making. However, the screening theory suggests that education mainly functions to prove to employers that someone has the skills and intelligence to complete a certain diploma; the diploma is used as a screening tool by employers. Individuals, then, should pursue a diploma that will make their skills most visible to employers, and employers will reward higher diplomas with higher wages. Wage information would be valuable according to this theory, but so would information about the relative probability of completing a diploma. Both the screening and human capital theories assume a "supple labour market" (p. 16), but, in practice, there is often a rigid wage structure and a "limited elasticity of demand and supply." Thus, information on the probability of finding a job given a certain type of education is very useful.

The labour queue theory assumes that that the wage structure is a given, and that knowledge and skills are not gained in the education system but by on-the-job training and experience. Employers choose workers who will cost the least to train, and educational attainment is only one variable in this decision, along with gender, age and work experience. Education is important in that it helps individuals get to the front of the queue. For proponents of this theory, wage information is not as relevant as information regarding the correlation of educational attainments and position in the queue. The logical consequence of this theory is that many people will pursue an education in order to get to the front of the queue, resulting in a large supply of over-educated workers who displace less educated workers, with no corresponding increase in wages. Information on the educational levels of people holding certain jobs, which can change over time, would be highly relevant.

Job matching theory suggests that an individual's productivity depends on the compatibility of the specific demands of a particular job and the abilities and knowledge of the person performing it. Workers are not in a position to predict their productivity until they are in a job and may be required to engage in a process of switching jobs until they find the right match. Wages vary according to one's position in the labour market. Thus, potential students need information about specific job characteristics, and the knowledge and skills required to give them an advantage in desirable jobs. To be relevant, labour market information would need to be categorized by occupation and education.

According to segmentation theory, the labour market is divided into two segments: a high-level, well-paid segment with upward mobility and secure employment, and a low-level,

poorly paid segment with horizontal mobility. The allocation mechanisms in each segment are different: high-level workers are more often well-educated males with good social backgrounds who do not belong to ethnic minorities. Low-level workers are more likely to be young, female, less educated and members of ethnic minorities. Thus, occupational information should be divided based on the size of the business and the branch of industry. Also, supply information should be differentiated by gender, age, educational level and ethnic minority status so that those in the lower segment can be traced and their position potentially ameliorated.

In all cases, the information provided would be based on forecasts because educational decisions must be based on future prospects, not the present situation. Heijke recommends several principles for labour market forecasting. First, he emphasizes that forecasts should be limited to the medium term. A five-year forecast is able to provide enough foresight for decision-making, but beyond this time frame, labour market processes and unforeseen circumstances compromise forecast accuracy. Forecasts should be repeated every two or three years, and should be evaluated and modified, if necessary. This will improve not only the accuracy of these forecasts but also the process of forecasting and, consequently, its reputation in terms of utility. Also forecasts should use "explanatory models" to account for the changing occupational and training structure of employment over time, instead of the usual fixed coefficients.

Heijke suggests that forecasts be "limited to giving a general characterization of the relation between demand and supply for broad categories of training over the whole forecast period" (p. 18). He cautions that quantitative information is perceived as more precise than it really is, and recommends that information for students should be in the form of general guidance regarding the prospects of different training categories, using qualitative descriptions and a general "good to poor" scale. The risks involved in particular choices of training should be mapped to provide information on the relative probability of finding a job. In addition, he stresses that "any understanding that already exists" about the education and labour markets should be used, including techniques of labour market analysis.

Heijke concludes by outlining the approach to forecasting used by the Dutch Research Centre for Education and the Labour Market (ROA). The ROA has incorporated many of Heijke's recommendations into their forecasts. It most closely follows the assumptions of the job matching theory, which Heijke believe is "more realistic" because it "makes allowances for the existence of disequilibria and sub-optimalities in the labour market" (p. 24).

Smith (2002)

Smith's (2002) article describes the COPS system, Canada's approach to labour market forecasting. The general structure of this model, which produces "anticipated imbalances by occupation and level of education" (p. 70), was described in the introduction. COPS uses two models, one for demand and one for supply. The demand side model uses macroeconomic forecasts to create a reference scenario for the overall growth of GDP. Using this reference scenario, a macroeconomic model is employed to create estimates of employment by industry for the forecast period, and these employment levels are then disaggregated into occupational categories. One of the advantages of the COPS model is that

it enables these coefficient estimates to vary over the forecast period, in contrast to the problematic "fixed coefficient" approach that is usually adopted. The demand for new workers can be determined by adding the replacement demand to the expansion demand, resulting in an estimate of the net number of jobs that will be available. Job prospects are presented in terms of both occupations (139 categories) and skill levels (five categories: managerial, professional, technical/paraprofessional/skilled, intermediate/semi-skilled, and labouring).

The supply side model is also very detailed, which mitigates the risk of inaccuracy created by the natural adjustment processes of the supply side of the labour market in response to demand signals. Supply includes school leavers, immigrants and re-entrants to the labour force. An occupation transition matrix developed by Statistics Canada estimates transition to occupation for graduates from high schools, community colleges, and universities, based the relationship between program curricula and specific occupations. One problem with this approach is that it ignores the "substitution" of skill sets (i.e., the application of skill sets to unanticipated occupations), which is a typical labour market response to shortages. The model does attempt to address common sources of error by estimating the number of students who will drop out and the number of immigrants who do not find jobs in areas consistent with their skills, education and stated occupational objective. The model does assume, however, that re-entrants will obtain work in the same occupation they occupied before leaving the labour force, which is another potential source of error.

The supply and demand models are integrated to arrive at medium-term estimates of excess supply, which is "defined as supply plus unemployed minus the number of new job openings" (p. 74). Excess supply may be negative, positive or nil. According to Smith, no regular assessment of forecast reliability is performed. One of the difficulties in performing such assessments is that there is "no accepted benchmark standard for accuracy" (p. 77). Also, to be useful, forecasts must be disaggregated by occupation, but accuracy appears to be compromised by higher levels of disaggregation. Because of the errors inherent in forecasting, and the natural adjustment process of the labour market, Smith stresses that COPS and other forecast approaches should be viewed as providing "conditional forecasts" (p. 78).

Smith maintains that labour market adjustment processes have three policy implications. First, government's key role is to provide information, which is a public good. Second, it is mainly the market itself that should, and will, react to labour market information, adjusting directly to those who demand its service. Third, shortages and surpluses are natural and should not be viewed as market failure; the goal is to facilitate the process of rapid and cost-effective adjustment.

Borghans and Willems (1998)

Borghans and Willems (1998) are also critical of forecasting approaches that focus on shortages and surpluses, and ignore the adjustment processes of the labour market through wages and substitution. Most forecasts assume constant relative wages and no substitution. Borghan and Willems argue that the forecasted gaps in supply and demand provide useful information if one employs "a labour market model that is completely cleared by wages" (p. 634).

The authors argue that policy-makers should not plan educational investments based on forecasted shortages and surpluses since forecasts are often inaccurate. Instead, policy-makers should provide information that enables students to adjust their educational investments to the needs of the labour market. Governments should also monitor the direction of the labour market so that they can intervene through investment in the educational infrastructure if necessary. These two goals require different information: for students, wage information is more important than supply and demand information, whereas for government, the focus is on "total quantities of supply or demand" on the labour market (p. 635).

The authors propose, but do not test, a model that takes into account these different needs. Labour market forecasts generally estimate a gap between supply and demand. Borghan and Willems posit that changes in wages will be proportionate to this gap: wages will go up relative to the degree of shortage, and down relative to the degree of surplus. Thus surpluses and shortages are good indicators of changes in wages, and vice versa, and consequently provide useful information. The sum of elasticities between supply and demand, which can vary widely, "determines the influence the gap might have on wage development" (p. 638). But a change in wages will alter the demand as well, so the authors provide an equation that estimates the distance from equilibrium of the supply forecast. This estimate shows policy-makers how adjustment processes may influence forecasted employment change. The authors conclude by stating that, by including the market adjustment processes in the model, they have solved the main problem associated with forecasting. It should be pointed out, however, that their model assumes that employers have unlimited flexibility in terms of adjusting wages, which is not generally the case. Moreover, since the authors do not apply or test their model, we must take their word regarding its utility and accuracy.

Castley (1996)

Castley (1996) suggests abandonning workforce forecasting and planning approaches, which, he argues, are too unreliable to guide enrolment decisions and investments. He presents a policy-focussed approach that determines the functions of labour market analysis (LMA) and provides a process for implementing LMA.

Castley lists the common criticisms of workforce forecasting: questionable assumptions about future GDP, sector growth and skills ratios; the failure to account for productivity changes; the difficulty of determining the skills or education needed for occupations; and the lack of adaptability to changing conditions. He further questions whether detailed long-term national projections are necessary or even useful. Most recruiting is done regionally, and national forecasts may not be relevant to specific regions. While policy-makers can open up more spaces in educational facilities, they cannot really plan to produce a specific number of people trained for a certain occupation. Forecasting approaches were originally used after WW II to create economic development through the training of a skilled workforce, but now there are other labour market issues, such as unemployment, migration, and workforce segregation by ethnicity and gender, which forecasting approaches cannot address.

Nonetheless, Castley acknowledges that there is still a need for some planning to ensure efficient and equitable functioning of labour markets. He advocates a labour market analysis approach, which he argues is more "flexible and responsive" (p. 16) and can address a wide

range of labour market issues from enrolment to unemployment. Such an approach would involve a fundamental shift: less emphasis would be placed on predicting occupational growth and more on measuring wages, conducting tracer studies and household surveys, creating educational profiles, addressing equity and poverty issues, and creating general training courses, among other things. However, he recognizes that LMA is a broad concept and "does not provide a satisfactory framework on which to base planning" (p. 17). He goes on to present a policy-focussed approach, which he hopes will provide this framework.

Castley argues that the role of a human resources (HR) analyst should be to oversee a "policy cycle," which begins with identification of a problem area that can be addressed by policy. The HR analyst will then need to engage in fact-finding and gathering data. Castley recommends that one central data bank be developed to house all data relevant to policy (it is important to exclude irrelevant, but available, data, which clogs the system). He further notes that data requirements need to be carefully specified when data need to be collected. The data must be analyzed and interpreted to be converted into useful information. The analyst's next task is to prepare a succinct, easily readable report, with specific goals for action and descriptions of various alternatives and possible outcomes that will enable rational choices to be made. The report should be given to interested parties and especially to key decision-makers. The policy is then implemented, which necessitates leadership by someone with a strong understanding of project identification and preparation. Finally, the implemented policy must be evaluated and feedback integrated into the next cycle. This approach underscores the importance of the HR analyst and provides a system to evaluate the results of the analyst's work in relation to the costs. Castley suggests that this is an improvement over the "ad hoc" nature of current workforce planning.

Willems (1996)

Willems (1996) addresses another issue associated with labour market forecasting: modelling replacement demand. He provides an overview of methods used to forecast replacement demand and to model labour supply in general. Willems pays particular attention to the workforce requirements approach, outlining the steps involved and the general mathematical equations used to obtain forecasts, and addressing the common criticisms. Like other authors we have discussed, he believes that the utility of forecasting has moved from educational planning to market transparency. Due to the lag time in educational adjustments, educational forecasting requires long-term projections, however because the quality and reliability of forecasting models decrease as the time period increases, Willems argues that there has been a change in the use of these estimates, mainly shorter projection times which provide information about labour market trends. Still, he emphasizes, these forecasts continue to be informative in terms of policy-making.

Willems goes on to outline ways of estimating the future supply of labour, especially future replacement demand. He distinguishes the three elements of the future labour supply, as outlined by Parnes (1962): the number of workers in the educational category of concern, the influx of new workers or re-entrants to the labour market, and the losses of workers in the market due to death, retirement or withdrawal. Estimations of future labour supply involve attempts to calculate outflow and inflow for each occupation in each year. On the outflow side, early retirement and movement from one occupation to another also need to be considered. On the inflow side, occupations that do not clearly link to type of education

will have to be modelled based on estimates of how many positions are filled by graduates with the generally expected educational profile and how many are filled by graduates with a different profile.

Willems discusses forecasting models in the United States, Canada, and western Europe, with a focus on the modelling of replacement demand. These models generally apply constant rates for immigration, death and retirement, differentiated by age category and gender, and neglect to model occupational mobility. Willems concludes that most of these models pay less attention to modelling replacement demand than to modelling labour demand (new jobs). He suggests that this may be due to the shift away from educational planning to transparency, or it may be a result of the "growing awareness that gaps between supply and demand are difficult to interpret," which has been discussed by other authors in our study. Another possible reason for the little attention paid to replacement demand is the lack of detailed data in many countries on the inflow and outflow of workers. Willems suggests that using "detailed stock data to estimate flows" is a promising means of making replacement forecasts.

Foot and Meltz (1992)

Foot and Meltz (1992) examine three occupational projections prepared in the 1960s and 70s: Meltz and Penz (1968), Ahamad (1969) and COFOR82 (1975). The authors first discuss the theoretical and methodological issues involved in creating these projections. The numerical *ex post* (projected versus realized) accuracy of the three projections is then examined, followed by a more detailed analysis of the fixed versus variable coefficient models used in one of the projections. The first theoretical issue the authors address is the means of modelling the future demand and supply for labour, since the projection is the anticipated net result of the two. The second theoretical issue is the potential for distortion that arises from projecting occupational requirements: by steering individuals into certain occupations, the projections themselves affect the operation of labour markets and the actual employment patterns. Foot and Meltz also address two methodological issues. First, two of the three projections they studied employ a different occupational classification than the one that was actually used in the target year. As a result, it is difficult to compare projections with realizations over time. The other methodological issue involves determining the characteristics of a "good" projection. Despite these issues, Foot and Meltz offer the following findings from their *ex post* analysis. First, they find that all three projections were within plus or minus 10% of the actual employment for the major occupational groups. However, the projection accuracy declines as the occupational classifications become more narrow and detailed. Their analysis also shows that the variable coefficient model is able to provide additional information for occupational projections.

Rosenthal (1999)

Rosenthal (1999) discusses the changes in the procedures used to develop projections over the past 50 years and presents data to determine if the quality of the projections improved as the projection procedures changed. From the mid-1940s to the mid-1960s, forecast information was presented solely in qualitative terms, although the projections

were based on statistical extrapolations or simple regressions. The use of an industry-occupation matrix in the 1964 to1975 projections represented a major change, but some occupations were not projected in the matrix due to data availability. From the mid-1960s to the late 1970s, three approaches were used to develop detailed industry projections: regression analysis, input-output analysis, and in-depth analyses of industries with poor statistical results in using past employment growth to indicate future growth. The 1978 to 1990 projections were developed using the first national industry-occupation matrix. Other major changes in the 1978 to 1990 projections were a focus on the input-output analysis and the abandonment of regression analysis procedures, which BLS economists believed would lead to more accurate predictions. Despite these changes, Rosenthal points out two constant factors. First, projections of the labour force were always a significant factor of the growth in total employment. Second, in-depth analysis of factors influencing changes in employment was always used in combination with statistical models. Rosenthal then analyzes in detail the accuracy and error in the projections by comparing the actual and the projected change in employment, and calculating the absolute average percentage error of the projections. He concludes that, although there were significant improvements in terms of the projection methodologies and data quality, the accuracy of the projections has not changed over time.

London Economics (2002)

The purpose of the London Economics (2002) report is to examine the use of regional forecasts and forecasting models by the Future Skills Wales (FSW) Partnership and to make suggestions on how to improve forecasting activities. After a discussion of the relevant literature, the report assesses data quality as well as the model quality. In addition, interviews with representatives from the FSW Partnership on the topic of regional forecasts are presented. There are quite a number of issues addressed in the report. Point estimates of forecasts of future GDP or employment growth in general are affected by a high degree of uncertainty. In Wales, particularly, the quality of regional data is often less than desirable, especially in the case of data from the sub-regional level, which limits the ability of information to be incorporated into the regional forecast models. Because of the data problem, the report is unable to fully assess and compare the quality of the three models it examines: the Oxford Economic forecast, Cambridge Econometrics and BSL. The three forecast models each have their own strengths and weaknesses; it is not possible to declare one superior to the other two.

3.3 Peer-Reviewed Empirical Studies

3.3.1 Introduction

The articles included in this section are empirical studies that use rigorous statistical or econometric methods to derive labour market forecasting models and test the performance of such models. Many propose forecasting models that differ from those used by government agencies, arguing that their models represent an improvement on the existing ones. The articles below apply a range of forecasting models to different areas of the labour market in various countries. The models may be applied to one small region of a country, many regions of a country, or a country as a whole. Unlike the descriptive, analytical and

critical articles, the empirical studies aim to create a forecasting model that produces superior forecasts. As a result, these empirical studies often focus on the derivation of the forecasting model and statistical tests of forecasting ability rather than the theoretical background of labour market needs. In the next subsection, the précis of the empirical studies are arranged in alphabetical order. A chart outlining the methodologies used in these studies is included.

3.3.2 Précis of Empirical Studies

Archambault (1999)

Archambault (1999) analyzes the Canadian Occupational Projection System (COPS) and compares it with his proposed projection equations. He specifies two criteria for good projections: they minimize the risk of introducing statistical bias, and they take economic information into account. Archambault argues that while the smoothing approach taken by COPS reduces the volatility of the occupational series, it does not enhance information quality concerning trends or the short-term behaviours of trends. This type of volatility originates from the excessive disaggregation of data into 67 industries and 139 occupations, and resulted in 9,313 series for projection. Archambault proposes a higher level of aggregation by reducing the number of industries from 67 to 17 and putting the 139 occupations into 25 groups so that the number of projections is reduced to 425. In addition, he proposes four projection equations to replace the smoothing approach. These four projection equations are linear with first difference specifications and differ on the basis of whether they have a trend component or a cyclical component. The criteria for evaluating the COPS projection and the projections of the four proposed equations are based on mean error projection, the Theil inequality coefficient, the bias that indicates the systematic projection error, the variance that indicates the model's ability to reproduce the variability of the series observed, the proportion of the covariance that measures the unsystematic error, and the part of error which cannot be attributed to the bias and the variance. The results indicate that the proposed method provides better projections than COPS at both the national and provincial level.

The intention of study is clearly stated. The two central research questions are clearly defined. The proposed method and the subsequent results provide a satisfactory response to the questions. The assumptions and the arguments are reasonable. The explanation of the design and the results are clear.

Bailey (1991)

Bailey (1991) analyzes trends, aggregate data, forecasts of occupational changes, and the changes in associated education level for certain occupations. He looks at aggregate U.S. data and trends over 15 periods from 1972 to 1988. The forecast is based on the projected growth rate from 1988 to 2000. First, Bailey analyzes the aggregate data on the number of jobs and education level, segregated by occupations. Then he uses simple regression by regressing the projected growth rate of the education level and finds that a 10% increase in the share of the occupation filled by workers with some college education is associated with a 3% increase in the projected growth of the occupation. Finally, he looks at the actual and projected growth rates and finds substantial gaps between the two for all occupations.

Although the intent of the study is clearly stated, Bailey's failure to deliver a clean and clear methodology results in an analysis of relatively poor quality. He uses the same reference for many of his arguments, so the objectivity of the study is affected. Second, the data tables are vague and ambiguous. In Table 3, there are two columns with the same heading but different numbers, and no explanation is provided. Bailey does not adequately explain or define the variables used in the study. Furthermore, he uses a single number to represent the growth of an occupation over 15 years. This may be misleading. There are numerous fluctuations within a 15-year span that cannot be captured by a single number. Bailey should make use of trend analysis and include some dispersion measure (e.g., standard deviation) of the growth rate. Bailey fails to acknowledge statistical significance in the study. All results in the study are reported without any statistical significance levels, so the claimed results may indeed be statistically insignificant. Finally, although he provides an explanation for the substantial gap between the actual and the projected growth rate of an occupation, the gap simply implies that the model fails to predict the growth of an occupation. Based on the above problems, as well as the limitations Bailey notes, the reliability of the results is highly questionable.

Bishop and Carter (1991b)

In their study of occupational projections, Bishop and Carter (1991b) argue with substantial evidence that Bureau of Labor Statistics (BLS) projections have systematically downward biased the growth of occupations that require higher education and training. The authors' regression analysis of trends in occupational shares using a logit model results in forecasts of much faster growth of higher-level occupations in 2000 than the growth projected by BLS. The authors identified the variables that had significant effects on occupational shares during the period from 1972 to 1989 and used them in the forecast model: a simple trend, the unemployment rate, the merchandise trade deficit as a proportion of GNP, and the ratio of personal computers used in business to total employment. Bishop and Carter then examine the supply and demand balance for college graduates by comparing past and projected percentage rates of change in employment in high-skill jobs (demand) to actual and projected rates of change in the stock of well-educated workers (supply). They predict that the labour market for college graduates will get tighter and the wage premium will continue to grow if supply cannot increase more rapidly.

The central research questions are clearly stated, and the arguments are strongly supported by the relevant statistics provided in the study. The analysis of the study is very clear and thorough. The data used are extensive and come from reliable sources. Issues regarding time series data and statistical significance in the forecast model are properly handled. However, the evidence may be too subjective and the findings, consequently, overstated. For example, the authors choose a ceiling of 20% instead of the more commonly used 25% ceiling because a 20% ceiling "fits the data slightly better." (p.238) A sound methodology for determining whether a model is useful should include estimating the lower bound or underestimating the model. If the lower bound still yields useful estimates, one can conclude the model is useful. Therefore, the authors should provide a forecast with a 25% ceiling as well and see if the resulting forecasts are robust to such a specification. Since the authors do not justify the model specifications, another concern is the robustness of other Datasets, such as data from other periods. The authors could have used the forecasting model to "predict" the past data and compare their predictions to the actual past data. Such a procedure can show the robustness of the forecasting model and provide a fair comparison

with the BLS historical projections. Furthermore, the authors point out the limitations and problems of BLS projections without mentioning the limitations of their own forecasting model. Given these concerns regarding robustness and overstated findings, this study is only of fair quality.

Blien and Tassinopoulos (2001)

Blien and Tassinopoulos (2001) created a forecasting method called ENTROP for labour market analysis and apply this method to the labour market in districts of western Germany. They claim that the superiority of this method lies in its flexibility (it allows linear inequalities) and efficiency (the computation time for the estimate is short). ENTROP first calculates autonomous regional trends and then estimates and transforms matrices of three dimensions: industry structure, the type of the region and the federal state that the district belongs to. Blien and Tassinopoulos then use the weighted mean squared error to examine the performance of the forecast and evaluate its reliability.

The intention of the study is clearly stated, and the methodology is consistent with the intention. However, the study is only of fair quality for the following reasons. First, the authors do not provide evidence or references for the choice of the three dimensions used in the ENTROP method. Furthermore, the reasons for using their method specification (the equations involved in estimating the forecasting matrix) are not explained. Without a solid theoretical background or reliable reference, the accuracy of the forecast might be coincidental and might not prove satisfactory for data from other periods. In addition, the authors do not explain why data from civil servants and low-income workers are excluded in the study. The data used in the study do not provide a true representation of the population of the districts in western Germany.

Byers and Peel (1994)

Byers and Peel (1994) look at series of macro-variables (output, real wages and employment) of the U.K. quarterly data from 1959 to 1991 and investigate the non-linearity of the data using the Brock, Dechert and Scheinkman statistics. Various regression estimations are made to adjust for the non-linearity based on the test results for different aspects of time series data. Tests performed in the study include the unit root test, the cointegration test, the Jarque-Bera test for normality and the ARCH test. Forecasts and relevant statistics from AR, SETAR, SECM and AECM models are compared. The results show that no one model outperforms all the others in all three data series. SETAR has smaller root mean squared error than AR. The forecasts of real wages and employment derived from AECM perform better than those derived from SECM. But the forecasts of output from SECM are superior.

The method used in the study is able to address the central research question. Assumptions are well justified by relevant references to other studies. The issues regarding time series data are well addressed by the various statistical tests. In addition, the estimation methods are decided according to the test results. The comparison of the forecasts is made objectively, based on relevant statistics. As a result, the model and results are of good quality. However, the authors do not specify the sources of the data. They also do not include any descriptive statistics about the three series. These issues affect the overall reliability of the study, so this study is only of fair quality.

Corvers and Heijke (2004)

Corvers and Heijke (2004) analyze the ROA labour market forecast. Their analysis focusses on two major features of the forecast: the gap approach and the substitution of education programs within occupations. The authors begin by explaining the theories behind these two features; they refer to the supply and demand of the labour market by education type. The gap approach provides insight into the size of the adjustment that is required to restore labour market equilibrium. The substitution demand takes into account the shifts in the occupational structure, upgrading and downgrading of occupations caused by technological developments, and shortages and surpluses in educational submarkets. Due to the insufficiency of information on the adjustment process, their empirical analysis focusses on illustrating the relevance of the substitution demand using the ROA labour market forecasts for the Netherlands for the period from 2003 to 2008. Their main finding is that the omission of the substitution demand in the forecasting model results in underestimated forecasts, especially for more highly educated workers.

The study does not involve any econometric or statistical techniques. The central research question is not clear, but the explanation of the theories behind the gap approach and substitution demand is useful. However, the empirical analysis on the actual ROA forecast is unable to provide support for the authors' theory of substitution demand. The analysis is based on some data that are not defined. The authors do not use any hypothesis testing. The data provided are unable to support the argument and conclusion.

Fauvel, Paquet and Zimmerman (1999)

Fauvel, Paquet and Zimmerman (1999) examine the short-term forecasting of employment variables (employment level, employment rate and unemployment rate) in Canada and its provinces. Based on an analysis of the results of relevant literature, the authors focus on the univariate ARIMA, multivariate VAR and indicator ARX models. The authors begin by discussing the principles of a good forecasting protocol: it follows a systematic path, can be easily replicated, and takes into account recent developments in the forecasting literature. Then the authors discuss the procedures and issues involved in the development of a forecasting protocol. These procedures include gathering and constructing variables, specifying the forecasting horizons, partitioning the Dataset, establishing a pre-forecasting estimation practice, and specifying the evaluation criteria of the forecasting performance. The actual application of the proposed forecasting protocol involves the assessment of the forecasting performance of the three selected models over one- to six-month horizons. These three models are reasonably good at forecasting the current context. The results show that, although the VAR model performs better in many cases, it does not surpass the other models in all short-term horizons. Furthermore, a model that surpasses the others in terms of forecasting accuracy may be itself surpassed in terms of forecasting direction.

The study is of good quality. The central research question is clearly stated, and the method proposed is able to address the question. The Dataset is clearly defined by the authors. The study is well supported by references to the literature. The choice of data and variables is also appropriate and objective. The examination of the forecasting performances of the selected model is thorough. The empirical results are reported in detail with rigorous use of

tables. The conclusion drawn by the authors is properly stated based on the empirical results. As a result, the findings of the study should be reliable and useful.

Hofler and Spector (1993)

Hofler and Spector (1993) employ a method that identifies various specifications for determining employment, as follows: equilibrium employment with flexible wages, demand-determined employment with fixed wages, and short-side determined employment with fixed wages. This method is much simpler computationally than former determination methods, which involve OLS estimations of a single equation derived from the standard macroeconomics aggregate labour market model and a series of hypothesis tests of the estimated results. The authors apply their method to data derived from the time series of U.S. annual employment and economic data from 1948 to 1984, including total labour force, male labour force, female labour force, real wages and GNP implicit price deflators. They find an insignificant estimated coefficient on real wage, which suggests that real wage is not a determination of employment. More importantly, the hypothesis testing results support the short-side determined employment model.

The study is of good quality and very clear in all aspects. The central research question is clearly stated. The procedures and steps taken in the study are well explained. These procedures lead to a well-defined conclusion that addresses the central research question. The data used in the study are obtained from reliable sources. The estimation equation is able to summarize and capture the essence of the standard labour market model. The authors take time series issues into consideration by applying appropriate statistical tests to detect any misspecification of the data. However, it should be noted that the authors use GNP implicit price deflators in the OLS estimation. GNP (the gross national product) captures the amount of final goods and services produced by citizens according to their place of residency. Citizens who reside overseas might have relatively little or no impact on the labour market. Since the labour market should address the employment conditions within the borders of the United States, the GNP implicit price deflator should be replaced by a better choice of variable.

Kolb and Stekler (1992)

Given the impossibility of conducting hypothesis tests of forecast accuracy, Kolb and Stekler (1992) evaluate the 1970, 1975, 1980 and 1985 BLS forecasts of employment in different industries using the author developed "information content" statistic. In each forecast, the percentage of the total employment for each industry is compared with the percentage of aggregate employment predicted for each industry to compute the statistics. After adjusting for the population size, the statistics follows the chi-square distribution so that it can be used for hypothesis testing on forecast accuracy. The results are mixed. The statistics with population size rounded to the nearest thousand are highly significant, but not the statistics rounded to the nearest ten thousand or hundred thousand.

Although the central research question is clearly stated, the reliability of the evaluation methodology proposed in the study is highly questionable. The results of different round-up choices are not consistent. The difference between the actual and the predicted values, instead of the differences between various round-up methods, should be a factor on the

evaluation results. Therefore, Kolb and Stekler's method is not appropriate for evaluating forecast accuracy.

Krolzig, Marcellino and Mizon (2002)

Krolzig, Marcellino and Mizon (2002) use a cointegrated vector autoregressive Markov-switching (MS) model to analyze the U.K. labour market in terms of four indicators: output, employment, labour supply and real earnings. With Markov-switching between the three regimes (recession, normal growth and high growth) built into the model, the proposed methodology provides better characterization of the changing phases of the U.K. economy than the alternative models (linear and non-linear VARs) used in other studies. The results show that standard linear models perform poorly and the MS-VECM model possesses satisfactory forecasting ability. The switches in the regimes are closely related to the changes in the U.K. business cycle.

The central research question is clearly stated. The methodology is able to address the question. The design of the study is rigorous and clean. The data sources are clearly documented. The authors clearly define and graph the variables. The results are discussed in term of statistical significance. The study is fairly objective because the authors use commonly used statistics to compare the results of the proposed model and alternative models used in other studies. Lastly, the conclusion is properly drawn based on the statistical results. Therefore, the study of Krolzig, Marcellino and Mizon is of very good quality.

LeSage (1990a)

LeSage (1990a) applies an error-correction mechanism method (ECM) to labour market forecasting. Such a method is well-suited to traditional export-based employment, with lower computational costs. The data consist of employment variables of durable, non-durable and non-manufacturing industries of eight metropolitan areas in Ohio. First, LeSage demonstrates the basis of the error-correction mechanism by showing the cointegration between decomposed export-based and local employment series in the long run using the Dickey-Fuller cointegration test. Then, a forecasting experiment is performed in order to compare the forecast performance of the proposed ECM model, the VAR model, the Minnesota-prior Bayesian VAR (MVAR), the block recursive VAR, and a mixed model of ECM and MVAR. The results show that both the percent root mean square error and the standard deviation of the forecast error are reduced by including the error-correction variable in the model.

The study is of good quality. The intention is clearly stated, and the design of the study is able to address the purpose. The methodology is well explained and consistent with relevant studies. The data are also clearly described. The measures for forecast accuracy are widely used in the literature. However, LeSage does not perform a test to determine whether the differences in RMSE and standard deviation are statistically significant. This might affect the reliability of the stated results.

LeSage (1990b)

LeSage's second study (1990b) closely resembles the one described above (LeSage 1990a): both studies propose the error-correction mechanism (ECM) method, test co-integration between variables, and conduct forecasting experiments on different models to compare their accuracy. In this study, the design is applied to the monthly labour market data of 50 Ohio industries from January 1976 to December 1985. The purpose of this study is to empirically test the theoretical claims of previous studies (i.e., Granger, 1986; Engle and Yoo, 1987) regarding the forecasting performance of ECM model. The results of Dickey-Fuller and Augmented Dickey-Fuller co-integration tests show that seven of the 50 industries exhibit co-integration between hours worked, earnings, and prices variable. An additional five of the 50 industries were possibly co-integrated. The experiments on the forecasting models (ECM, VAR, MVAR, BVAR and BECM) are performed separately, based on the degree of co-integration: co-integrated, possibly co-integrated and non-Co-integrated. Forecasting accuracy is measured based on the average mean absolute percent errors between one- to 12-month horizons. The results show that the ECM model produces forecasts with smaller margins of error than the alternative VAR or BVAR models for the co-integrated industries, which is consistent with the previous studies. However, contrary to the claim of Engle and Yoo, the BECM model performs well in the longer forecast horizons for both co-integrated and non co-integrated industries.

The study is of good quality. The intention is clearly stated, and the design is able to address the purpose. The methodology is well explained and consistent with the relevant studies. The data are clearly described. The measure for forecast accuracy is widely used in the literature. However, LeSage does not perform tests to determine whether the differences in average MAPE are statistically significant. This might affect the reliability of the stated results.

Li and Dorfman (1995)

Li and Dorfman (1995) propose a composite forecasting model, created by combining individual autoregressive leading indicator (ARLI) models, to forecast state-level employment changes . The authors argue that the composite model should outperform individual ARLI models due to its adaptability to changes in economic conditions. The authors use four leading indicators in all possible combinations of exclusion and inclusions, yielding 16 ARLI models. These leading indicators are average weekly manufacturing wages (MFG), the initial new unemployment claims (IUC), the new housing permits issued (NHP) and an index of help wanted advertising from the City of Atlanta (HAD). The output variable is the growth rate of Georgia's non-agricultural employment. A dichotomous dependent variable is constructed to represent correct or incorrect *ex post* forecasts for each model. This dependent variable is then regressed in a logit model to yield the probability of its being correct in the next forecast. These probabilities are normalized and used as time-varying average weights. These time-varying average weights are able to reduce the impact of lesser performing models through logit regression. The authors' claims for their model, however, are not justified: the logistic regression result shows that one of the two individual ARLI models outperformed the composite model.

Although the data used in this study are of good quality, the study itself is only fair. The authors' explanation of the methodology is sometimes unclear. The value of some important parameters, such as loss structure c1 and c2, are chosen without any justification. More importantly, one is uncertain about the usefulness of the proposed method. We are often interested not only on the direction of change (upward or downward) in employment, but also the size of the change. The proposed model only focusses on the direction of the change. It is unlikely this method can produce forecasts of the change in the size of related revenue collection by the state government as the authors claim.

Longhi *et al.* (2005)

Longhi *et al.* (2005) apply the artificial neural networks (ANNs) technique to forecast regional employment. Unlike regression techniques, ANNs allow the number of parameters estimated to exceed the number of explanatory variables. However, it is difficult to interpret the relationship between dependent and explanatory variables. The empirical evidence is obtained by applying various forecasting models to the panel data of employment variables from 327 regions of western Germany. These models have the same set of explanatory variables, but their means of estimation vary. The forecast performance of three ANNs models, two maximum likelihood (ML) estimation models with random effects, and two models that combine ANNs and ML are compared using statistical indicators commonly used in the time-series literature. These indicators include mean absolute error, mean absolute percentage error, and mean square error and its components, which provide different measurements of the difference between the actual employment and the predicted employment. Using Theil statistics, these seven forecasts are each compared with the forecast of a simple naïve model used by the German authority. The resulting Theil statistics show that all seven forecasts outperform the forecast of the simple model. Although forecasts of several models combined often perform better than forecasts of a single model in the literature, one of the ANNs models performed best in this study.

The Dataset was well documented by the authors. The methods were clearly explained. In addition, the results were reported in a very clear manner. The use of statistical indicators is rigorous, but the authors should have considered providing the statistical indicators of the forecast of the simple naïve model as they did for the other forecasts. Although the central research question is clearly stated and the study design fits the question, it is not clear whether the ANNs could be used more generally to forecast regional employment effectively in different contexts. Such a technique is highly arbitrary; there is no consensus on the number of iterations required for deciding weight computation and the "hidden units." It is, in certain respects, a trial and error process until the desirable pattern emerges. This might be computationally impractical. Moreover, the best specification identified in the current study is unlikely to be robust given another set of data because there is no theory to support such ad hoc estimation. In other words, ANNs might be best suited to the current data.

Longhi and Nijkamp (2005)

In their study of labour market forecasting, Longhi and Nijkamp (2005) incorporate spatial information using panel techniques in their models for regional labour market forecasts. The authors demonstrate spatial autocorrelation. They propose various models that differ in their estimation methods (fixed effect or maximum likelihood) and the ways in which the spatial

information is incorporated (in a lagged variable or in the error term). These models are applied to data from 326 districts in western Germany for *ex post* forecast comparison. The authors found that spatial error models outperform the non-spatial models in general. However, the spatial error models do not outperform the non-spatial models and the naïve no-change model in some cases. The authors argue that incorporating spatial information is useful for regional forecasts models, but the means of incorporation plays a significant role.

Despite a few typos, the study is of good quality. The central research question is clearly stated. The explanation of the model equations is clear as well. Although descriptive statistics are not provided, the data source and the variables are clearly stated. The assumptions made throughout the studies are reasonable. The authors use various measures, rather than one simple measure, to assess the forecasting accuracy of the different models.

Magura (1998)

Magura's (1998) study is an extension of previous studies on the use of input-output information as a Bayesian prior in forecasting employment. In this study, Magura investigates the forecast accuracy of methods using spatial information, as well as methods combining input-output and spatial information. The forecast accuracy of previous models (UVAR, BVAR and IOBVAR) is compared with the two extended models proposed: spatial BVAR (SBVAR) and input-output spatial BVAR (IOSBVAR). The employment forecast models are constructed for four industries (primary metal products, fabricated metal products, machinery, and transportation equipment) in five mid-western states (Indiana, Michigan, Ohio, Kentucky and Pennsylvania). The forecast accuracy is evaluated based on the number of the lowest root mean squared error (RMSE) values for each of the five models for various forecast horizons. The results indicate that the IOSBVAR model, which uses input-output weighted inter-industry relationships in combination with spatial information, has the greatest number of the lowest RMSE values of the five models, which suggests its superiority as a forecasting method.

The intention of the study is clearly indicated. Its design allows it to answer the central research question. The construction of the forecasting models is well explained. The ideas presented are consistent with those in the literature and are supported by various references to other studies. The nature of the data used is clearly documented, but descriptive statistics are not provided. Although RMSE is a typical measure for forecast accuracy, the way that Magura uses this measure is problematic. His conclusion is based on the number of the lowest RMSE values for each of the five models for various forecast horizons. That is, for every particular forecast horizon, Magura compares the five RMSE values obtained from the five models. The actual differences between these RMSE values in each comparison are not clear. It is also unclear if these differences are statistically significant. For example, a model might have low but insignificant RMSE values for many of the horizons and significantly large RMSE values for other horizons. Such a model could have the highest number of the lowest RMSE values. Yet, it may not have superior performance. To provide a more objective evaluation, Magura should assess the forecast accuracy of the models with more than one measure, as well as reporting the results statistics.

Partridge and Rickman (1998)

In their study of regional employment forecast, Partridge and Rickman (1998) apply the Bayesian vector autoregression approach to forecasts of regional employment in the state of Georgia. Their approach is an improvement on those of previous studies for various reasons. They incorporate regional input-output coefficients rather than national coefficients. Their model also includes final-demand effects on inter-industry relationships, and links to national and world economies. Based on two sets of out-of-sample forecasts, the forecast accuracy of the two models adopted by the authors, IOVVAR and IOMVAR, are examined and compared with the forecasts of three other models, AR, UVAR and MVAR, which were used in related studies. Mean absolute percent errors are used as the measure of accuracy. No one model provides superior forecasting accuracy in all cases. In other words, the results fail to show that IOVVAR and IOMVAR can forecast Georgia's employment with comparative accuracy. The authors discuss the relative performance of the models with respect to the length of forecast, the type of industry and the economic scenario. AR and MVAR models are more accurate in the short run, but UVAR and IOMVAR are more accurate in the long run. In addition, UVAR and IOMVAR models are more accurate when forecasting in those sectors that are most reliant on the local economy. These two models also provide more accurate forecasts in scenarios of fluctuating economies.

The central research question is clearly formulated, and the study design is able to address the question. The execution of the study is rigorous and clean. The derivation and estimation of the models are discussed in great detail. In addition, the other models used in related literature are included and discussed in the study. Although the value for the parameters in the models was chosen to maximize the accuracy of the forecasts for all models, the authors provide appropriate sensitivity tests to check the effects of different parameter values. The authors also use appropriate statistics to verify the model specifications. They try to ensure that different models are comparable by taking logarithms and first difference, so that all coefficients are approximately equal to elasticities. As a result, this study is of good quality.

Patuelli *et al.* (2003)

In the study by Patuelli *et al.* (2003), the multicriteria analysis (MCA) technique is used to evaluate the forecast performance of the neural network (NN) models and of the NN models extended with the genetic algorithm (GA) in the context of given conditions. This empirical analysis is performed on German labour market data. The NN methods are "statistical goodness-of-fit techniques based on learning principles, where, through repetitive experiments of individual data, a hidden structure is identified." (p.6) The GA is able to update information on the underlying population in each iteration of the algorithm. MCA was developed to enable systematic evaluations; it allows the search of a dominant alternative, given different results from different evaluation criteria. The eight evaluation criteria considered in the study can be grouped into four categories. Mean squared error, mean absolute error and mean absolute percentage error are used to ensure statistical relevance. The number of epochs and number of NN weights account for computational efficiency. Stability and generalizability are used as reliability indicators. The inclusion of a wage variable facilitates interpretation. The authors conclude that the MCA technique is

more useful in evaluating NN models than individual statistical tests. When one or two models appear to dominate the other models, MCA is able to confirm this dominance.

Although the central research question is clearly stated, it is not clear whether the study fully addresses the question. The purpose of the study is to investigate whether the proposed technique, MCA, is useful in evaluating the forecasting performance of NN models. The MCA is only able to provide relative rankings for the NN models considered in the study. Unlike standard statistical tests that offer clear cut-offs for statistical significance of forecast accuracy, the MSA does not provide such cut-offs. As a result, it is possible that the NN model with the highest ranking has poor forecast accuracy but performs better than the other NN models.

Puri and Soydemir (2000)

In their study of employment forecasting, Puri and Soydemir (2000) use Bayesian vector autoregressive (BVAR) models to forecast industrial employment in southern California. Their models incorporate both state and national variables in order to capture the economic interactions between the region and the nation. State variables include the total employment for five counties (Los Angeles, Orange, Riverside, San Bernardino and Ventura) and the employment in the three largest industries (manufacturing, retail and services). National variables include a leading indicator composite index and an industrial production index. The out-of-sample forecasts obtained from the BVAR models are compared to the forecasts obtained from unrestricted VAR and best-fit ARIMA models. The root mean squared errors (RMSE) and the Theil's U statistics are used to compare the forecasting accuracy of these models. The results show that the BVAR models perform better than VAR and ARIMA models, especially for longer horizons. Among BVAR models, models with a looser prior produce more accurate forecasts than those with a tighter prior.

Puri and Soydemir's study is of good quality with some minor weaknesses. The central research question is clearly formulated, and the study design is able to address the question. The execution of the study is rigorous and clean. The study is filled with relevant references from the literature. The authors try to ensure an objective comparison by using best-fit ARIMA as the benchmark, which has the highest forecasting power. However, some of variables and the data sources used are not clearly explained. In addition, the authors do not test whether the differences in RMSE and the U statistics for different models are statistically significant. As a result, it is not clear whether the model with the smallest RMSE or U statistics produces the most accurate forecast.

Rickman (2001)

In his study of Oklahoman labour market forecasting, Rickman (2001) uses Bayesian methods to incorporate regional input-output information into the regional employment matrix of the econometric model. Based on the IMPLAN modelling system, the forecasting equations incorporate input-output inter-industry, local final demand, domestic export and foreign export linkages by sector into single variables of intermediate demand, local final demand, domestic export demand and foreign export demand by aggregation. The out-of-sample forecast accuracy of the proposed Bayesian models (ECBAS, MECBAS, REBIO and REBIOX) are compared to some traditional forecast approaches (AR1, REMIO and

UNIO). Accuracy is measured by mean absolute percentage errors (MAPEs) and relevant statistics of MAPEs. In general, the results indicate that there is no model that surpasses all the others under all circumstances. On average, however, Bayesian models yield the lowest MAPE across all employment sectors and forecast horizons studied. In addition, the IO integrated Bayesian models are more accurate than the classic AR1 model in the long run.

This study is of good quality. The central research question is clearly formulated, and the study design is able to address the question. The execution of the study is rigorous and clean. The derivation and estimation of the models are discussed in great depth. In addition, the other models used in the related literature are included and discussed in the study. Although the value for parameters was chosen to maximize the accuracy of the forecasts for all models, the authors provide appropriate sensitivity tests to check the effects of different parameters. The authors also use appropriate statistics to verify the model specifications.

Rickman and Miller (2002)

Rickman and Miller (2002) evaluate various approaches to incorporating inter-industry relationships into the forecasting equations of regional industry employment. The approaches to constructing the models can be differentiated in four ways. First, the approach uses either input-output (IO) information to select which industry employment variables to include or IO-based aggregate demand variables. Second, approaches differ in the degree of endogeneity between industry employment levels. Third, they differ according to which model selection procedure (IO information, stepwise or Bayesian) is used to specify the independent variables. Fourth, they differ according to whether they place restrictions on the estimated coefficients or not. These models are applied to the regional data of Oklahoma for empirical evidence. Evaluations are made based on the mean absolute percent errors and the employment multipliers. Given the data on the Oklahoman economy, models with restrictions on inter-industry relationships through IO linkage in aggregation yield more useful employment multipliers. But the resulting forecasts are similar to those of models for selected individual industries. The Bayesian model averaging selection procedure does not generally produce superior forecasts.

The central research question is clearly formulated, and the assumptions made by the authors are appropriate. The explanation of the various approaches is fairly thorough and clear. The execution of the study is rigorous and transparent. The data used are clearly documented. Selected descriptive statistics are provided. The results reported address the purpose of the study. Appropriate statistical tests and treatments for non-stationarity are used in the study. Therefore, this study is of very high quality.

Sarantis and Swales (1999)

In their study of labour market forecasting, Sarantis and Swales (1999) use four models (a time-varying parameter model, a regression model, a state space model and an ARIMA model) to forecast regional employment growth in the service sector of Great Britain. The in-sample and out-of-sample performances of the four models and a professional structural model known as NIERC-MRM forecasts are compared. In-sample results indicate that the four models have satisfactory performance. The out-of-sample performance of the models

varies with respect to regions and forecast horizons. Significantly, the NIERC-MRM model did not perform better than the four proposed models.

This study is of good quality in all aspects. First, the data source is clearly documented. Although descriptive statistics are not reported, the authors provide a graph to illustrate the trend of the employment growth variable. The authors explain the construction of the models in a clear manner. The assumptions made are reasonable and supported by references. The forecasting accuracy of the models is evaluated in various dimensions using a range of statistical indicators. Appropriate statistical tests are performed. The results are discussed based on statistical significance. The implications of the results are objectively stated according to the results.

Sweeney (2004)

In his study of regional labour market forecasting, Sweeney (2004) proposes a demographically driven labour supply approach in contrast to the commonly used demand requirement approach. The demand requirement approach assumes a perfectly elastic labour supply, which would not act as a constraint on demand. Sweeney argues that the elasticity of labour supply depends on the region and occupation. The demographically driven labour supply approach (or the demography approach) focusses on long-term migration and mobility trajectories of interacting state occupational labour markets, as well as the information from age-sex disaggregated flows. The estimation equation is based on a system of balance equations determining regional occupational and demographic variables. Due to the data quality, Sweeney is unable to draw a convincing case for the superiority of his model. However, he argues that neither of the models, if considered alone, would be ideal for long-term forecasts.

The central research question is clearly stated. The theory used to support the argument is also convincing. The derivation of the system of balance equations is rigorous. However, the representation of the forecast results is very unclear; it is difficult to refer the findings to the relevant data provided in the tables. The data used in the study are not appropriate for the model proposed; the model would require very detailed data. Using the indirect estimation method to generate the data may alter their true characteristics. Sweeney should have used an alternative estimation model that scopes with the available data. As a result of these weaknesses, the study is not able to provide any insightful conclusions.

Trivez and Mur (1999)

Trivez and Mur (1999) develop the transfer functions model to forecast sectoral employment in the short run. Their study assumes that the region of interest is sufficiently small and its economic structure is close to that of the nation. As a result, the authors are able to ignore the endogeneity between regional and national variables and hypothesize a unidirectional flow from the national to the regional level. The series of national employment variables of a given sector are used as the inputs for forecasting the regional employment variable of that sector. Unlike other regional forecast studies, the authors take into account the outliers in the given time series data during the estimation of the transfer functions model. The statistics used in the *ex ante* diagnostic checking stage confirm that the residual from the estimated model follows a Gaussian white noise process, which is a

requirement for the transfer functions model. *Ex post* diagnostic checking, measured by guaranteed prediction error (GPE) and relative real prediction error (RPE), also confirms the precision, accuracy and corroboration of the estimated model. In addition, the statistics from the estimated model are compared with the corresponding statistics from an ARMA model. Although, when measured by mean RPE, the forecasts of the two models show no substantial differences in terms of accuracy, all forecasts from the ARIMA model yield higher mean MGPEs, which suggests the higher information content contained in the forecasts from the transfer functions model. The authors used the estimated model to predict the sectoral employment of Aragon in 1996.

Trivez and Mur's study is of good quality in all aspects. The data are obtained from a prominent, clearly documented source. Although the authors do not present descriptive statistics for the data, they do provide graphs to illustrate the data series. In terms of method quality, the study design is able to address the central research question. The execution of the study is clearly explained, and the arguments and assumptions are often supported by relevant studies. In addition, the authors use a variety of techniques to assess the quality of the proposed model, which enhances the objectivity of the study. Statistics and tests are used appropriately in the study. However, one should be aware of the limited applicability of the proposed model. It can only forecast the employment in small regions that have an economic structure that is very similar to that of the nation.

Turner, Wallis and Whitley (1994)

Turner, Wallis and Whitley (1994) examine the important differences in the simulation properties of four leading U.K. macroeconometric models based on the features of labour market specification. More specifically, they investigate and resolve the discrepancies resulting from three simulations. The models studied are the quarterly models from the London Business School (LBS), the National Institute for Economic and Social Research (NIESR), Her Majesty's Treasury (HMT) and the Bank of England (BE). The first set of simulations involves a change in government expenditure. The simulation result shows that the employment effects are consistently much larger for the NISER model than for the other models, possibly due to the unusual representation of employment in the behavioural equation, estimated by output and a time trend, of non-manufacturing industry. Supported by rigorous statistical tests, the authors suggest a new specification of this equation, which induces more dynamics (more flexible to changes in variables) and replaces output with actual output, to remove this discrepancy. The second set of simulations involves three standard policy simulations (government expenditure, income tax and employer national insurance contributions) on the four models to examine the relationship between changes in unemployment and employment. Despite an expectation of a one-to-one ratio between the fall in unemployment and the increase in employment, the fall in unemployment observed in the BE and NIESR models is only a small proportion (converges to about 20%) of the increase in employment, in contrast to the other two models (where it converges to over 90%). The authors show the poor long-run properties of the NIESR models with standard diagnostic tests. In addition, they point out the specification and data problems associated with the BE model. Replacing the unemployment behavioural equations in the BE and NIESR models with the alternative participation equation proposed by the authors effectively narrows the differences between the models. The third set of simulations involves an income tax cut. The simulation results of the HMT and BE models show decreases in the

level of nominal and real earnings, which is contrary to the increase one would expect. The authors explain that this discrepancy is due to the inclusion of the retention ratio (ratio between net and gross wages) in the wage equations of these two models. The discrepancy can be narrowed by excluding the retention ratio in wage equations.

The central research question is clearly stated, and the design of the study is able to address this question. The rigorous statistical tests used in the study are appropriate and suitable for comparing the simulation results and addressing the discrepancies arisen from the simulation. The incorporation of graphs and tables for presenting results enhances the clarity of the study. All simulation and estimation results are clearly reported by the authors. The authors' suggestions for improving the existing models are justified and supported by reasonable assumptions and appropriate statistical tests. However, the descriptions of data are minimal, which affects the quality of the study.

van Eijs and Borghans (1996)

van Eijs and Borghans (1996) investigate the optimal aggregation level of the RAS model for manpower forecasting. They estimate and evaluate the performance of nine variants of the RAS model, from total segregation into clusters by each industry and each education level to total aggregation of a single cluster, using data on the Dutch economy. Two hypotheses are proposed and tested. The competitiveness hypothesis assumes wages and wage changes are uniform over the labour market and are not industry specific. The technology hypothesis assumes production functions are industry specific. Conclusions are drawn based on the relevant statistics derived from the sum of relative quadratic forecast errors. Both the competitiveness and technology hypotheses cannot be rejected, which means that applying RAS to the Dutch economy as a whole yields better forecasts than applying it to partial clustering of industries.

The model used in this study is of very high quality. The authors deliver very clear explanations and derivations of the standard RAS model and its variants. They are able to provide sensible microeconomic interpretations of the RAS model in the current context. Their assumptions are well supported by relevant studies. Despite the high quality of its model, this study is only of fair quality overall because of its data, which affect the objectivity of the results. The data sources are not reported in the study. Furthermore, no descriptive statistics or definitions of variables are provided. As the authors themselves point out, the quality of data is extremely important when assessing the forecasting performance. There is insufficient information on the data provided in this study. The finding that total aggregation of the RAS yields the best performance might due to the noise at a lower aggregation level, a form of poor data. It is possible that RAS of low aggregation would perform better when it is applied to another Dataset of better quality. In order to provide more conclusive findings, the authors should include relevant statistics or historical trends in order to assess data quality.

West (2003)

West (2003) examines whether the absolute and relative accuracy of *ex ante* forecasts can be predicted. Panel data on regional employment forecasts of 19 MSAs of Florida from regional structural econometric (RSEMS) models, regional stochastic time-series (RSTIMS) models, and the difference between the two (DIFER) are used. In addition to this Dataset, regional characteristics and period effects are incorporated into a simple fixed-effect panel model. The results show that for RSEMS and RSTIMS forecasts the estimated model's ability to predict forecast accuracy increases with the step length of forecast. In addition, the model performs better when forecasting at low step lengths. The results also indicate that, for all three sets of models, period effects become more consistent and systematic as step length increases.

The data used in the study are obtained from reliable sources. Descriptive statistics are provided. West also follows appropriate procedures when dealing with stationarity, simultaneity and multi-colinearity. Although West clearly demonstrates the shortfalls of the previous studies in the literature, the methodologies used in the study are not very clearly explained. The author should provide an explanation of the construction of the RSEMS and RSTIMS models. It is uncertain which regional economic characteristics are used in which model and how the two models differ.

Wong, Chan and Chiang (2005)

In their study of the construction labour market in Hong Kong, Wong, Chan and Chiang (2005) use the Box-Jenkins approach to develop ARIMA models for forecasting five construction labour market indicators: employment level, productivity, unemployment rate, underemployment rate and real wage. The quarterly time series data used in this study were obtained from the General Household Survey and the Survey of Construction Outputs conducted by the Census and Statistics Department of the HKSAR government. These surveys provide data from the first quarter of 1983 to the first quarter of 2004 with a total of 85 data points. The first 80 data points (from the first quarter of 1983 to the fourth quarter of 2002) are used for the in-sample estimation of the forecasting model, while the other five data points (from the first quarter of 2003 to the first quarter of 2004) are compared to the corresponding out-of-sample forecasts to evaluate predictive adequacy. The mean absolute percentage error and the Theil's U statistics are the measures used in the evaluation. The results indicate that the model proposed has good forecasting ability (except in the area of employment level) with regard to the construction labour market.

The study is of very good quality. The central research question is clearly stated. The statistical methods used in the study are appropriate and suitable for addressing the research question. Furthermore, appropriate statistical tests are performed and reported. The quality of the data is also high. The data come from reliable sources and cover 99% of the population. However, there are two minor issues that are worth noting. First, it is uncertain whether the variable used in calculating productivity—construction output in dollars— is normalized or deflated to a common year's value. Second, the authors do not provide any descriptive statistics or trends for the time series data of the indicators.

3.3.3 Characteristics of Methodologies of Empirical Studies

The table below outlines and describes the data, populations, model specifications and measures of forecasting for each of the empirical studies analyzed in this review. Readers will notice, when comparing the studies in the table, how diverse the research and model specifications are in the area of labour shortage forecasting.

Characteristics of Methodologies of Empirical Studies

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Archambault	1999	LFS data	The 25 2-digit occupational groups in the National Occupational Classification (NOC)	Estimated from 1984 to 1996 Projected for 1997 and 1998	Canada, at national and provincial levels	Ordinary least squares (OLS) regression	Mean error projection U-Theil Bias that indicates the systematic projection error (BP) Variance that indicates the model's ability to reproduce the variability of the series observed (VP) Proportion of the covariance that measures the unsystematic error, and that which cannot be attributed to the bias and the variance (CP)

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Bailey	1991	1989 Bureau of Labor Statistics (BLS) projections 1988 Current Population Survey (CPS) Silvestri and Lukasiewics, 1989	Aggregate data on occupational and educational changes, segregated by occupations	Estimated from 1972 to 1988 Projected for 1988 to 2000	United States, at the national level	Statistical analysis Simple regression	Ratio of projected to actual rate
Bishop & Carter	1991b	BLS projections CPS OES Workforce 2000 U.S. Bureau of Labor Statistics	Aggregate data, segregated by occupations Change in employment in high-skill jobs Change in the stock of well-educated workers	Estimated from 1972 to 1989 Projected for growth rates from 1988 to 2000	United States, at the national level	Logit	N/A
Blien & Tassinopoulos	2001	Employment statistics of western Germany	Cross-sections of all people employed on 30 June in any of the periods between 1987 and 1997	Estimated from 1997 data Targeted 1999	25 districts of western Germany	ENTROP method	Mean squared error (MSE)

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Byers & Peel	1994	N/A	130 seasonally adjusted quarterly observations on measures of real wages, employment and output	Estimated from second quarter of 1959 to fourth quarter of 1988 Projected for first quarter of 1989 to third quarter of 1991	U.K., at the national level	Cointegrating regressions by Phillips-Hansen procedure Autoregression of order one (AR(1)) SETAR SECM AECM	Root mean squared error (RMSE)
Corvers & Heijke	2004	Research Centre for Education and the Labour Market (ROA) of Maastricht University	Labour market forecasts by ROA from 2003 to 2008	N/A	The Netherlands, at the national level	N/A	N/A

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Fauvel, Paquet & Zimmerman	1999	Human Resources Development Canada Statistics Canada Manpower Inc.	First set of series: monthly data, seasonally adjusted, from the first two sources Second set of series: quarterly survey of employers conducted by Manpower Inc.	Short-term, one- to six-months, between January 1983 and September 1998	Canada, at national and provincial levels	ARIMA VAR ARX	RMSE Mean absolute error (MAE) Mean absolute percent error (MAPE) BP VP CP U-Theil
Hofler & Spector	1993	The Statistical Abstract of the United States Economic Report of the President (1987)	Time series of annual employment and economic data such as total labour force, male labour force, female labour force, real wages, and GNP implicit price deflators	Estimated from 1948 to 1984	United States, at the national level	OLS regression Log-linear	Hypothesis testing
Kolb & Stekler	1992	Bureau of Labor Statistics (BLS), US Department of Labor	1970, 1975, 1980 and 1985 BLS forecasts	N/A	United States, at the national level	N/A	Information content statistics

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Krolzig, Marcellino & Mizen	2002	*Economic Trends* and *Monthly Digest of Statistics,* published by the U.K. Office of National Statistics	Seasonally adjusted quarterly data series of wages and salaries, employment, workforce, gross value added (GVA), implicit deflator of GVA, and unemployment benefit claimants	Estimated from the fourth quarter of 1965 to the first quarter of 1991 to produce forecasts for the second quarter of 1991 to the first quarter of 2001 Estimated from the fourth quarter of 1965 to the first quarter of 1998 to produce forecasts for the second quarter of 1998 to the first quarter of 2001	U.K., at the national level	MSIH-VECM MSIH-VAR VAR	Root mean square prediction error Mean absolute prediction error
LeSage	1990a	Ohio Bureau of Employment Services, Labor Market Information Division Employment and Earnings monthly reports of the Bureau of Labor Statistics	Monthly magnitudes of local and export employment Three industry categories of durable, non-durable and non-manufacturing	Estimated from February1971 to December 1982 and forecast over 1983 and 1985	Eight metropolitan areas of Ohio	ECM VAR MVAR BVAR BECM	Percentage root mean squared error (PRMSE) Standard deviation of forecast errors

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
LeSage	1990b	*Labor Market Review,* publication of the Labor Market Division of the Ohio Bureau of Employment Services	Monthly labour market data (employment, hours and average hourly earnings) for fifty industries at the two- and three-digit SIC levels of aggregation National consumer price index for all items	Estimated from February 1977 to December 1982 Forecasted for 1983 to 1985	Ohio, at the state level	ECM VAR MVAR BVAR BECM	MAPE
Li & Dorfman	1995	Georgia Department of Labor U.S Department of Commerce The Conference Board	Monthly data (all variables except HAD were seasonally adjusted and converted to annual rates of growth)	Estimated from January 1968 to January 1981 Forecasted for February 1981 to December 1991	Georgia, at the state level	Autoregressive leading indicator model by OLS regression Logit model by maximum likelihood estimation	Number and percentage of most accurate forecasts

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Longhi *et al.*	2005	German Institute for Employment Research	Panel data Employment data collected from individual firms Variables: number of full-time workers and mean regional daily wages	Estimated from 1987 to 1997 and forecasted for the growth rate between 1998 and 1999	327 regions in western Germany	Artificial neural networks (ANNs) Maximum likelihood with random effect	MAE MAPE RMSE MSE BP VP CP U-Theil
Longhi & Nijkamp	2005	German Institute for Employment Research (IAB)	Panel data Employment data collected from individual firms Number of full-time workers employed and mean regional daily wages	Estimated from 1987 to 1999 Forecasted for 2000, 2001 and 2002	326 regions in western Germany	Non-spatial models Spatial models	MAE MAPE MSE BP VP CP

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Magura	1998	Bureau of Labor Statistics	Monthly data of two-digit industry employment figures	Estimated from 1972 to 1993 Forecasted for 1994	Five mid-western states of the United States	UVAR BVAR IOBVAR SBVAR IOSBVAR	Number of the lowest RMSE values
Partridge & Rickman	1998	The Georgia Department of Labor *Employment and Earnings* *Federal Reserve Bulletin* IMPLAN regional input-output system	Monthly establishment employment data covering January 1972 to June 1993 U.S. employment data Foreign exchange rates IMPLAN(a non-survey, industry-based, input-output model with levels of disaggregation for 528 sectors)	Estimated from January 1972 to December 1991 and forecasted for January 1992 to June 1993 Estimated from January 1972 to June 1990 and forecasted for July 1990 to June 1993	United States, at the national level	Bayesian vector autoregression (BVAR) Unrestricted vector autoregression (UVAR) Autoregressive model (AR)	Number of most accurate forecasts MAPE

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Patuelli *et al.*	2003	Institute for Employment Research, Nuremberg, Germany	Panel data on German regional labour market conditions	Estimated from 1987 to 1999 Forecast for 2000 and 2001	327 districts of western Germany in nine economic regions	Neural network (NN) methods NN method with genetic algorithm	Statistical measures (MSE, MAE, and MAPE) Computational measures (number of epochs and number of weights) Reliability measures (stability and generalizability indicators)
Puri & Soydemir	2000	Bureau of Labor Statistics	Quarterly data of employment of the three largest industries (manufacturing, retail and services) for five counties (Los Angeles, Orange, Riverside, San Bernardino and Ventura) in the southern California region	Estimated from the first quarter of 1983 to the third quarter of 1994 Forecasted for the fourth quarter of 1994 to the fourth quarter of 1996	Southern California region	BVAR ARIMA VAR	RMSE U-Theil

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Rickman	2001	Bureau of Labor Statistics Minnesota IMPLAN Group, 1996 Massachusetts Institute of Social and Economic Research	30 industries that comprise the total private non-farm employment in Oklahoma Quarterly, non-seasonally-adjusted employment data from the second quarter of 1983 to the third quarter of 1998	Estimated from the second quarter of 1983 to the fourth quarter of 1994 Forecasted for the first quarter of 1995 to the third quarter of 1998	Oklahoma, at the state level	AR(1) ECBAS MECBAS UNIO REMIO REBIO REBIOX	MAPE Statistical significance of MAPE differences Number of most accurate forecasts
Rickman & Miller	2002	IMPLAN Bureau of Labor Statistics Massachusetts Institute of Social and Economic Research	30 industries that comprise the total private non-agricultural employment Quarterly non-seasonally-adjusted employment data from the second quarter of 1983 to the third quarter of 1998	Estimated from the first quarter of 1985 to the fourth quarter of 1994 Forecasted for the first quarter of 1995 to the third quarter of 1998	Oklahoma, at the state level	OLS OLS with stochastic restrictions Input-output linkage selection Stepwise regression Bayesian model selection	Weighted MAPE MAPE

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Sarantis & Swales	1999	Department of Employment's surveys of employers, Great Britain	Quarterly series, seasonally unadjusted employment	Estimated from the third quarter of 1978 to the fourth quarter of 1991 Forecasted for the first quarter of 1992 to the fourth quarter of 1993	U.K., at the regional level	Time-varying parameter model (TVP) Regression model (MVM) State space regression model ARIMA	MAE RMSE Turning point accuracy (TPA)
Sweeney	2004	1990 Census Public Use Microdata Sample, US Bureau of the Census March 1985 and January 1987 Current Population Survey, US Bureau of the Census	The complete Dataset needed for the model is "beyond the reach of any existing secondary data source." Sweeney uses indirect estimation methods to recover the data required from multiple sources.	N/A	U.S. regions	System of balance equations	MAPE Mean error
Trivez & Mur	1999	Labour Force Survey by National Institute of Statistics of Spain	Infra-annual time series on sectoral employment and regional labour force Quarterly data on sectoral employment	Estimated from the third quarter of 1976 to the fourth quarter of 1995 Forecasted for the four quarters of 1996	Aragon, a Spanish region	Transfer function model (stochastic time series) Maximum likelihood estimation	Guaranteed prediction error (GPE) Real prediction error (RPE)

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Turner, Wallace & Whitley	1994	Her Majesty's Treasury (HMT)	Time-series data that are relevant to the labour market Variables used in the four leading macroeconometrics models from LBS, NIESR, HMT and BE	Estimated from the third quarter of 1972 to the second quarter of 1986	U.K., at the national level	Simulations IV regression OLS regression	N/A
Van Eijs & Borghans	1996	N/A	Panel data From total segregation into clusters by each industry and each education level to total aggregation of a single cluster	Three estimation horizons: 1979 to 1981, 1981 to 1983, and 1983 to 1985	The Netherlands, at the national level	RAS Hypothesis testing	Sum of the relative quadratic forecast errors
West	2003	Unemployment insurance reports (ES 202 data) U.S. Census Bureau	Panel data on regional employment forecasts from RSEMS and RSTIMS	Estimated from the fourth quarter of 1985 to the second quarter of 1992	19 metropolitan statistical areas of Florida	SUR Fixed effects	MAPE

Author	Year	Data source	Data description	Estimation/ forecasting horizons	Labour market(s) under consideration	Model specification(s) & estimation method(s)	Measure(s) of forecasting performance
Wong, Chan & Chiang	2005	Quarterly General Household Survey (GHS) Quarterly Survey of Construction Outputs	Time series data on employed persons, gross construction output, unemployment rates and underemployment rates	Estimated from the first quarter of 1983 to the fourth quarter of 1992 Forecasted for the first quarter of 1993 to the first quarter of 2004	Construction industry in Hong Kong	ARIMA model Maximum likelihood estimation	Variance of forecast error MAPE U-Theil

4.0 QUALITY ASSESSMENTS

4.1 Introduction

In this review, we evaluate the quality of an econometric study according to three aspects: the quality of data, the quality of the model and the quality of the result. We further categorized the following characteristics of each aspect. For data quality, there are four categories: data source, data completeness, representative sample and data description. For model quality, there are also four categories: type of analysis, model assumptions, model sepcification and choice of variables. Finally, for result quality, there are three categories: statistical significance, estimation bias and objectivity of the discussion. Although some of these characteristics are related to each other, we believe each of them is essential for producing a good study with meaningful results. Therefore, the same scoring weight is assigned to all characteristics. A score of three indicates the study possesses good quality, a score of two, fair quality, and a score of one, poor quality. In order search for studies with the highest qualities, the scoring is downward biased. That is, if a study were between fair and poor for a particular characteristic, it would be considered poor.

4.1.1 Criteria for assessing the quality of data

Facet of Study:	Study scored 1 if:	Study scored 2 if:	Study scored 3 if:
Quality of Data:			
Data source	The data are obtained from surveys/questionnaires conducted by the researcher. All sources of data are not documented.	The Dataset is obtained from a single external source. Some sources of data are not documented	The Dataset is obtained from multiple external sources. The sources of all data used in the study are clearly documented.
Data completeness	A substantial amount of data is missing. The missing data seriously affect the study results.	The researcher provides a reasonable explanation for missing data. The missing data are not important to the study. The missing data do not seriously affect the study results.	There are no missing data.
Representative sample	The chosen sample is a poor representation of the population of interest.	It is uncertain whether the chosen sample could serve as a good representation of the population of interest.	The chosen sample serves as a good representation of the population of interest.

Facet of Study:	Study scored 1 if:	Study scored 2 if:	Study scored 3 if:
Data description	The researcher does not describe the unit or the definition of the variables.	The unit or definition of the variables is described but not clear.	The unit and the definition of the variables are clearly described.
Quality of Model			
Type of analysis	The study does not employ any econometric methods and relies entirely on qualitative methods (trend analysis and correlation analysis).	The study uses only econometric methods for estimating the results.	The study is a mix of quantitative and qualitative analyses. The researcher mainly uses econometric methods for estimating the results. Some qualitative analyses are provided for enhancement.
Model assumptions	Assumptions are unreasonable. Assumptions are made without any explanation.	Assumptions are not relevant to the study. Assumptions are non-intuitive. The explanation by the researcher is not very convincing.	Assumptions are intuitive. Assumptions are used in other relevant studies. Assumptions are necessary and important for the study, and the researcher has provided reasonable explanations.

Facet of Study:	Study scored 1 if:	Study scored 2 if:	Study scored 3 if:
Model specification	The model is unable to address the central research question. The specification is uncommon, and the researcher does not provide any statistical test. The specification is uncommon, and the researcher either does not provide any explanation or provides a poor explanation. The chosen specification does not account for the issues arising from the type of data used.	The model is able to address the central research question. Although the researcher does not justify or test the specification, it is common in relevant studies. The specification is consistent with the type of data used by the researcher.	The model appropriately addresses the central research question. The researcher tests the validity of the functional form specification. The researcher justifies the specification with reliable references. The specification is well suited to the type of data used by the researcher.

Facet of Study:	Study scored 1 if:	Study scored 2 if:	Study scored 3 if:
Choice of variables	The model does not include many of the influential factors. There are no control variables. Proxy variables, if any, are not relevant to their underlying factors. Instrumental variables, if any, are weak.	The model includes many of the influential factors. Some control variables are missing. Proxy variables, if any, are relevant to their underlying factors. Instrumental variables, if any, are adequate.	The model includes all of the influential factors. Proper control variables are used. Proxy variables, if any, are highly relevant to their underlying factors. Instrumental variables, if any, are strong.
Quality of Results			
Statistical significance	Estimates that capture statistical significance are not reported. Results are not discussed in terms of statistical significance.	Estimates that capture statistical significance are reported, but the researcher does not discuss the results in terms of statistical significance.	Estimates that capture statistical significance are reported. Results are discussed in terms of statistical significance.
Estimation bias	The results are biased.	The results may be biased, but the direction of the effects is reliable.	The results are unbiased.

Facet of Study:	Study scored 1 if:	Study scored 2 if:	Study scored 3 if:
Objectivity of the discussion	The researcher discusses the results in a subjective manner. Implications and inferences are made that are beyond the scope of the estimated results. The discussion substantially overstates the estimated results.	The discussion slightly overstates the estimated results.	The researcher discusses the results in an objective manner. Implications and inferences are made on the basis of the estimated results.

4.2 Overall quality

4.2.1 Introduction

With 11 criteria, the possible scores range from 11 to 33. A study with a score of 28 or above is considered to be a good study. A good study possesses high quality in all three aspects: data, model and result. The stated findings of a good study should be objective and reliable. A study with a score between 22 and 27 is considered to be a fair study. A fair study only possesses good quality in some of the criteria we seek. There are some minor methodological flaws in a fair study, but the results still provide some insight into the research question. The stated findings should be interpreted with caution. A study with a score of 21 or below is considered to be a poor study. Since there are major methodological flaws in a poor study, the stated findings are unreliable. Readers of this report should note, however, that since the criteria is designed for evaluating econometrics studies, studies that do not involve econometrics techniques would generally receive scores in the poor range. The details regarding the methodological quality of the studies are included in the précis in Section 3.4.2. The following chart lists the scores obtained by the empirical studies for each characteristic.

4.2.2 Chart of methodological quality assessment

*Non-econometrics studies

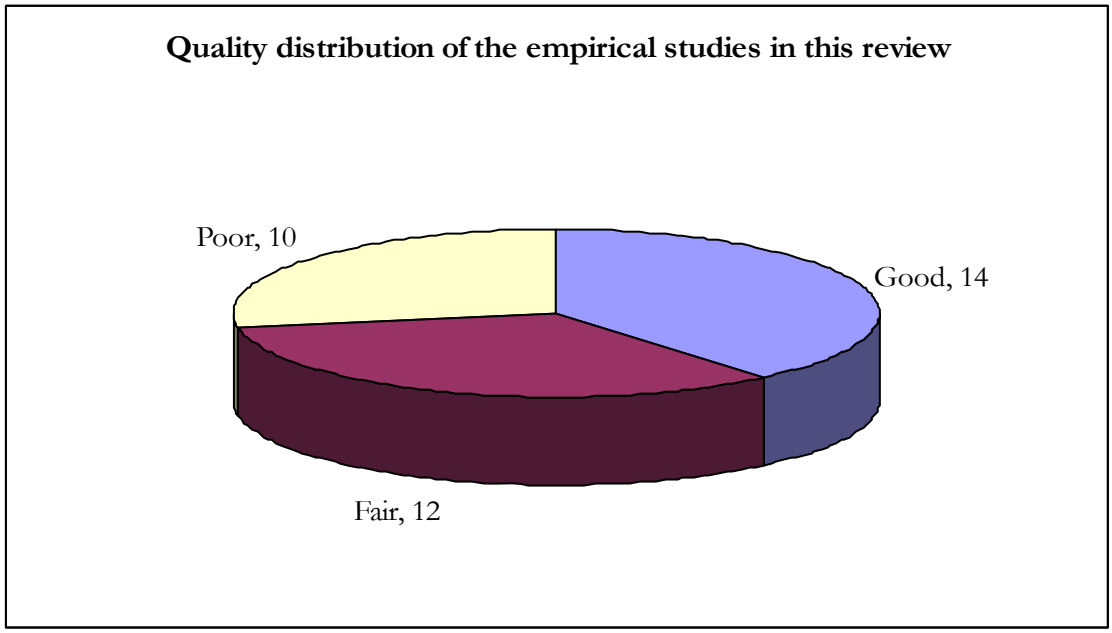
Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
Archambault (1999)	2	3	3	3	3	3	3	3	3	3	3	**32**	**Good**
Bailey (1991)	3	3	3	1	1	1	1	1	1	2	2	**19**	**Poor**
Bishop & Carter (1991b)	3	3	3	2	2	2	2	2	3	2	2	**26**	**Fair**
Blien & Tassinopoulos (1994)	2	2	1	2	3	2	2	2	3	3	2	**24**	**Fair**
Borghans & Willems (1998)*	3	2	1	1	1	1	1	1	1	2	2	**16**	**Poor**
Byers & Peel (1994)	1	2	2	2	2	3	3	3	3	3	3	**27**	**Fair**
Castley (1996)*	1	1	1	1	1	1	1	1	1	1	1	**11**	**Poor**
Corvers & Heijke (2004)*	2	1	1	1	2	2	2	3	1	1	1	**17**	**Poor**

Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
Fauvel, Paquet & Zimmerman (1999)	3	3	2	3	3	3	3	3	3	3	3	**32**	**Good**
Foot & Meltz (1992)*	3	3	3	2	1	1	1	1	1	2	2	**20**	**Poor**
Heijke (1996)*	2	1	1	1	1	1	1	1	1	1	1	**12**	**Poor**
Hofler & Spector (1993)	3	3	3	2	2	3	3	2	3	3	3	**30**	**Good**
Kolb & Stekler (1992)*	3	3	1	3	1	2	1	1	3	2	2	**22**	**Fair**
Krolzig, Marcellino and Mizen (2002)	3	3	3	3	3	2	3	3	3	3	3	**32**	**Good**
LeSage (1990a)	3	2	3	2	2	2	3	2	2	3	3	**27**	**Fair**
LeSage (1990b)	3	2	3	2	2	2	3	2	2	3	3	**28**	**Good**
Li & Dorfman (1995)	3	3	3	2	3	2	2	2	1	2	2	**25**	**Fair**
London Economics (2002)	3	3	3	3	1	1	1	1	3	3	3	**25**	**Fair**

Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
Longhi *et al.* (2005)	3	3	3	2	2	2	2	3	3	3	2	**28**	**Good**
Longhi & Nijkamp (2005)	3	3	3	2	2	3	3	3	3	3	2	**30**	**Good**
Magura (1998)	3	3	2	2	2	3	3	3	1	2	2	**26**	**Fair**
Partridge & Rickman (1998)	3	3	3	2	2	3	3	3	3	2	3	**30**	**Good**
Patuelli *et al.* (2003)	3	3	3	2	2	1	2	3	2	2	2	**25**	**Fair**
Puri & Soydemir (2000)	2	3	2	2	3	3	3	3	2	3	3	**29**	**Good**
Rickman (2001)	3	3	3	2	2	3	3	3	3	3	3	**31**	**Good**
Rickman & Miller (2002)	3	3	3	2	2	3	3	3	3	3	3	**31**	**Good**
Rosenthal (1999)*	2	2	3	1	1	1	1	1	1	2	2	**17**	**Poor**
Sarantis & Swales (1999)	3	3	3	2	3	3	3	3	3	3	3	**32**	**Good**

Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
Smith (2002)*	3	1	1	1	1	1	1	1	1	1	3	**15**	**Poor**
Sweeney (2004)*	3	1	2	2	3	2	2	3	2	3	3	**26**	**Fair**
Trivez & Mur (1999)	3	3	3	2	3	2	3	3	3	3	3	**31**	**Good**
Turner, Wallace & Whitley (1994)	2	2	2	2	3	3	3	3	3	3	3	**29**	**Good**
Van Eijs & Borghans (1996)	1	1	1	1	3	3	3	3	3	2	2	**23**	**Fair**
West (2003)	3	3	3	2	2	2	2	2	3	3	3	**28**	**Good**
Willems (1996)	3	3	2	3	1	1	1	1	1	1	1	**18**	**Poor**
Wong, Chan and Chiang (2004)	3	3	3	2	3	3	3	2	3	3	3	**31**	**Good**

4.2.3 Pie chart of the quality distribution of the empirical studies in this review



5.0 RESULT

5.1 Summary of forecast results

This section focusses on the results of the empirical studies reviewed here. These studies generally test existing or newly proposed models of forecasting; they use statistical techniques, which permit more robust conclusions.

In the following table, a list of models from each empirical study and the results regarding their forecasting ability are presented. Proposed models are models originating in the study itself. Alternative models are not derived from the study; rather, they are models that are compared to the proposed model or existing models to which new data are applied.

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
Archambault	1999	Projection equation with a trend component and a cyclical component Projection equation with a trend component but no cyclical component Projection equation without a trend component but with a cyclical component Projection equation without both trend and cyclical components	COPS	Proposed models have better forecasting performance than COPS at both national and provincial levels
Bailey	1991	Simple regression: regressing projected growth rate on education level	None	Proposed model fails to predict the growth of occupation Substantial gap between the actual and the projected growth rate of occupation

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
Bishop & Carter	1991b	Logit model	BLS projection	Unknown, both models project the future rather than *ex post*
Blien & Tassinopoulos	2001	ENTROP	N/A	Reliable model
Byers & Peel	1994	N/A	AR(1) SETAR SECM AECM	No one model outperformed all the others in all three data series
Corvers & Heijke	2004	N/A	N/A	N/A
Fauvel, Paquet & Zimmerman	1999	N/A	ARIMA VAR ARX	VAR models have better performance in many cases, but they do not always dominate the other models over all short-term horizons
Hofler & Spector	1993	Linear model with ordinary least squares	N/A	Supports the short side determined employment model rather than demand determined employment model
Kolb & Stekler	1992	N/A	1970, 1975, 1980 and 1985 BLS projections	The results are mixed, depending on round-off method
Krolzig, Marcellino & Mizon	2002	Cointegrated vector autoregressive Markov-switching model	Linear and non-linear VARs	Proposed model possesses satisfactory forecasting performance Linear models are strongly rejected

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
LeSage	1990a	ECM (Error Correction model) BECM (A mixed model of ECM and MVAR)	VAR MVAR (Minnesota-prior Bayesian VAR) Block recursive VAR	Proposed models produce better forecasts than alternative models Including the error-correction variable in the model reduces forecast errors
LeSage	1990b	ECM BECM	VAR MVAR Block recursive VAR	ECM model produces forecasts with smaller margins of error than the alternative models for the cointegrated industries BECM model performs well in the longer forecast horizons for both cointegrated and non-cointegrated industries
Li & Dorfman	1995	ARLI (autoregressive leading indicator) models Composite model	N/A	One of the ARLI models outperforms the composite model, which is contrary to the authors' hypothesis
Longhi *et al.*	2005	3 ANNs models 2 models of ML (maximum likelihood) estimation with random effects 2 models that combine ANNs and ML	Simple naïve no-change model	All proposed models outperform the forecast of the simple model One of the ANNs model has the best performance in this study

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
Longhi & Nijkamp	2005	Model without spatial autocorrelation, using FE (fixed effect) estimator Model without spatial autocorrelation, using ML estimator Model without spatial autocorrelation, using FE estimator, allowing some spatial heterogeneity in regression coefficients Model without spatial autocorrelation, using ML estimator, allowing some spatial heterogeneity in regression coefficients Spatial lag model with FE estimator Spatial lag model with ML estimator Spatial error model with FE estimator Spatial error model with ML estimator	Simple naïve no-change model	On average, the maximum likelihood estimations seem to perform better than fixed effects estimations for models Spatial lag model do not outperform the non-spatial models Spatial error models outperform other models Only spatial error models outperform the naïve no-change model in almost all cases
Magura	1998	SBVAR IOSBVAR	UVAR BVAR (Bayesian VAR) IOBVAR	IOSBVAR model has the greatest number of lowest RMSE among the five models, suggesting its forecast superiority

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
Partridge & Rickman	1998	IOVVAR IOMVAR	AR UVAR MVAR	AR and MVAR models are more accurate in the short run UVAR and IOMVAR are more accurate in the long run UVAR and IOMVAR models are more accurate in forecasting those sectors that are most reliant on the local economy
Patuelli *et al.*	2003	NN (neural network) GA (NN models extended with the Genetic Algorithm in the context of conditions)	N/A	MCA technique is useful beyond judging by individual statistical tests
Puri & Soydemir	2000	BVAR	UVAR ARIMA	BVAR models perform better than VAR and ARIMA models, especially for longer horizons A BVAR models with a looser prior produces more accurate forecasts than one with a tighter prior.
Rickman	2001	ECBAS MECBAS REBIO REBIOX	AR(1) REMIO UNIO	No one model dominates the others under all circumstances

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
Rickman & Miller	2002	Models differed in four ways: 1. according to whether they use input-output (IO) information to select which industry employment variables to include or IO-based aggregate demand variables 2 .according to the degree of endogeneity between industry employment levels 3. according to whether model selection procedures are used to specify the independent variables or not 4. according to whether restrictions are placed on the estimated coefficients or not	N/A	The forecasts of models with restrictions on inter-industry are similar to models with selected individual industries The Bayesian model averaging selection procedure does not generally produce superior forecasts
Sarantis & Swales	1999	Time-varying parameter model Regression model State space model ARIMA	NIERC-MRM	In-sample results indicate all four proposed models have satisfactory performance NIERC-MRM model does not perform better than the four proposed models Out-of-sample performance of the models varies with respect to regions and forecast horizons
Sweeney*	2004	Model with demographically driven labour supply	Model with labour demand requirement	Due to data quality, the author is unable to draw convincing conclusions on which model is preferred
Trivez & Mur	1999	Transfer function model	ARIMA	Forecasts from the ARIMA model yield higher mean MGPEs, which suggests that higher informative content is contained in the forecasts from the transfer functions model

Author	Year	Proposed model(s)	Alternative model(s)	Result(s) on forecasting ability
Turner, Wallis & Whitley	1994	N/A	Model from London Business School Model from the National Institute for Economic and Social Research Model from Her Majesty's Treasury Model from the Bank of England	No conclusions on forecasting ability because results are based on simulations rather than forecasting
Van Eijs & Borghans	1996	RAS models with different aggregation levels	N/A	Applying RAS to the Dutch economy as a whole yields better forecasts than applying it to partial clustering of industries
West	2003	Regional structural econometric (RSEMS) models Regional stochastic time-series (RSTIMS) models The difference between RSEMS and RSTIMS (DIFER)	N/A	Forecast accuracy increases with step length of forecast Time effects become more consistent and systematic as step length increases
Wong, Chan & Chiang	2004	ARIMA models	N/A	Proposed model has good forecasting ability for the construction labour market in Hong Kong

5.1.1 Overall results of empirical studies

Is it possible to accurately forecast labour market needs? This review suggests several approaches to this central research question. Many of the studies are able to confirm the forecasting accuracy of their proposed models. Some demonstrate their superiority to existing models, including the models used by government agencies (e.g., COPS in Canada and NIERC-MRM in the U.K.). However, the types of forecasting model used are quite diverse and therefore not easily comparable. For instance, many incorporate regional-specific factors. None of these studies use data from the same region and the same forecasting horizon, so we are unable to make convincing comparisons of their results. Also, it is not clear if the models would perform as well in another forecasting horizon. Furthermore, the kinds and number of forecast accuracy measures used are very different across the studies, and there is no consensus on the relative merits of these measures. Lastly, the forecasting models focus on different aspects of the labour market. For example, some try to forecast the change in the overall employment rate, while others attempt to predict the change of employment in specific industries.

As a result, we are unable to draw definite conclusions about the best forecasting models for labour markets in general. In addition, we cannot confidently infer that one particular kind of forecasting model consistently outperforms the others. The studies suggest that there is no single forecasting model that can accurately forecast labour market needs in all situations. While some of the proposed models show an impressive level of accuracy in forecasting the labour market in a particular market, without controlled replicability, the consistency of their forecast accuracy remains uncertain.

The unique setting of each study also makes it difficult to distinguish a superior forecasting model. However, we are able to highlight the features of the studies with the most robust methodologies. These are studies which scored the highest according to our scheme of quality assessment in Section 4. They fulfill the criteria for a good labour market forecasting model.

Archambault's (1999) study is one of the four that were considered of high quality. He uses previous research results and logical arguments to isolate the weaknesses of the existing model used by the government. Furthermore, he constructs new models and adjusts the available data to overcome the weaknesses he detected. This process enriches the existing literature, and it gives the government the opportunity to consider an improved model, which could replace the one currently in use. Archambault is able to demonstrate the superiority of the adapted models by applying all the models to the Dataset. Moreover, because there is no consensus on which measures should be used to evaluate forecast accuracy, the fact that Archambault uses more than one measure also makes the study more objective.

Krolzig, Marcellino and Mizen (2002) also compare the forecasting accuracy of the proposed models with that of alternative models using various measures. An additional feature, which enhances the quality of this study, is that the authors apply the models to different forecasting horizons in order to show the consistency of the forecasting results. This makes

the results even more convincing. The data used in the study is of good quality. The Dataset is well defined, and the authors provide clear and documented definitions of the variables.

Fauvel, Paquet and Zimmerman (1999) do not propose any forecasting model; they evaluate the forecasting accuracy of existing models. This study is comparable to the other highly regarded studies because it is clear and its in-depth analysis of the existing models relies on theoretical support and empirical evidence. After a thorough discussion of the models using evidence from the existing literature, the authors apply the models to the same Dataset in order to make fair comparisons. As is the case in other studies of a robust nature, the models are tested over different forecasting horizons, using various measures for forecasting accuracy. Similar to the study of Fauvel, Paquet and Zimmerman, Sarantis and Swales (1999) also uses various measures to make comparison on the forecasting accuracy of the proposed and alternative model described in their study. The conclusion and findings of this study are objectively stated according to the estimated results.

Ultimately, it appears that some forecasting models have the ability to estimate labour needs in very specific circumstances. What remains unknown is whether it is possible to develop a single model that will accurately forecast in a range of situations, under various conditions.

6.0 CONCLUSION AND FURTHER CONSIDERATIONS

The descriptive, analytical and critical studies reviewed provide extensive theoretical background on the determination of labour market needs. The most important issue raised is the impact of forecasting on existing and future labour markets. Most agree that good projections have the potential to enhance labour market efficiency. Conversely, poor projections can provide incorrect information for labour market participants. Some authors criticize the weaknesses of the existing labour market projections or forecasting models, while others are more constructive, providing suggestions for improving forecasting methods and thereby their contribution to labour market efficiency. Whether these suggestions would improve labour market projections can only be determined by the implementation of rigorous empirical studies.

The literature on labour market forecasting models is highly empirical in nature, so it is often difficult to explain the model specification theoretically. The studies in the forecasting literature often rely on statistical tests of model validity. The few studies that construct structural models using economic theories often provide inferior forecast results. While the results of higher quality empirical studies can suggest whether or not a proposed model can yield accurate forecasts, one should not conclude that the model provides explanations for underlying causal relationships. A likely reason for the poor forecast results of structural models may be their rigid, inflexible nature.

Forecasting results are strongly influenced by the quality of the data. However, it is often difficult to assess the data quality. Studies do not usually include a discussion of the data quality or of the availability of data suitable for the proposed forecast model. Furthermore, the authors often focus their efforts on deriving the forecasting models and pay much less attention to the data. The appropriate level of aggregation or disaggregation of employment data is often an issue. If a poor Dataset is applied to a model, poor forecast results may wrongly be ascribed to a poor forecasting model. It would be beneficial if future authors included more discussion of the data quality. The robustness and application of the forecasting models are also of concern. The proposed models are often region specific. The assessment of the forecast accuracy is often made on the basis of a particular region over a particular period. It is uncertain if the results could withstand other model settings. As a result, repeat assessments over different forecast horizons and in different regions should be required prior to any assertion regarding the robustness of a given model.

While evaluating the various articles for quality, rigorous methodology, reliability and transferability to other sectors or regions is difficult enough, determining whether concrete conclusions can be drawn from these disparate studies requires forbearance. Dismissing forecasting models as valueless because they are unable to predict changes or cannot adequately represent all the variables inherent in a complex system like labour markets seems as impetuous as abandonning a nation's entire economic policy because of the results of one labour forecast. Forecast models are designed to be tools used in conjunction with other information to help decision-makers better understand present and future needs. Thus, it is more prudent to assess what constitutes a useful forecast than to make blanket statements about the possibility of using forecasting models to predict labour market needs.

As noted, high quality data are extremely important when conducting forecasting research. Readers should be apprised of sample sizes, data-gathering methods, the origin of the data (country/region) and the occupations that the data assesses. Understanding the scope of the data used in a model may provide clues to the model's relevance. For instance, a model based on national data may include occupations of little or no relevance to a particular region. Conversely, a regional model may reflect the specific needs of a particular locale while offering little of value to decision-makers in a different region or at the national level. While it may seem obvious, it should be stressed that overestimating the value of available data for a given purpose jeopardizes the results before other analysis has begun.

In addition to assessing the quality of the data, determining whether or not the forecasting model is suitable for its intended use is imperative if meaningful analysis is to be performed. As Campbell (1997a) points out, the strengths and weaknesses of various models must be understood so that mistaken assumptions are minimized. Such assumptions often arise when models used to forecast at one level are applied to other levels. Campbell observes that sectoral-level forecasts are occasionally used to make occupational forecasts in manpower projections. Any inference based on the conversion of results from one area of inquiry to another does not seem a sound basis for extracting significant results. Additionally, the evidence suggests that the longer the duration of the projection, the more likely it is to become inaccurate and unreliable. Other issues that should be considered when assessing whether a given model is appropriate include the following: the location and the agent of the data collection, the number of variables represented, the comprehensibility of the results for non-specialists and the inclusion of uncontrollable factors. While it may be unrealistic to assume that all uncontrollable factors can be accounted for in any given forecast, those forecasts that make no reference to any problems they attempted to rectify should be approached with caution.

GLOSSARY

AR model

Autoregressive model

ARX model

A class of model that uses multivariate systems of dynamic equations. ARX models can be seen as an extension of AR and VAR models.

ARCH

Autoregressive conditional heteroskedasticity

AECM model

Asymmetric error correction mechanism model

ARIMA

Autoregressive integrated moving-average

Bias

The difference between the expected value and the population parameter value of an estimator

Clustering

A technique for data analysis that involves partitioning a Dataset into subsets whose elements share common traits

Cointegrated

If two or more series are themselves non-stationary but a linear combination of them is stationary, the series are said to be cointegrated.

Endogeneity

A term used to describe the presence of an endogenous explanatory variable, which correlates with the error term

ENTROP

A forecasting method that is based on a special entropy-optimizing procedure

Ex ante

Beforehand

First difference

The change of variables is used instead of the level of the variables. It is often done due to non-stationarity of the disturbance term or time-trends in the variables.

Hypothesis testing

A statistical test of the null or maintained hypothesis against an alternative hypothesis

Logit model

A model for binary response where the response probability is the logit function evaluated at a linear function of the explanatory variables

Maximum likelihood estimation

A broadly applicable estimation method where the parameter estimates are chosen to maximize the log-likelihood function.

Mean error projection

The root of the mean forecasting error, which measures deviations in the projected series from the observed series

Mean square error

The square of the difference between the actual value and the predicted value

MS-VECM

Markov-switching vector equilibrium correction model

SECM model

Symmetric error correction mechanism model

SETAR model

Self-exciting threshold autoregressive model

Standard deviation

The root mean square (RMS) deviation of the values from their arithmetic mean. It is the most common measure of statistical dispersion; it measures how widely spread the values in a Dataset are.

Theil inequality coefficient

A measure of the degree to which one time series differs from another

VAR model

Vector-autoregressive model

REFERENCES

Studies Included for Analysis

Archambault, R. (1999). *News COPS occupational projection methodology* (Technical No. T-99-3E). Hull, Quebec, Canada: Human Resources Development Canada. (Applied Research Branch)

Bailey, T. (1991). Jobs of the future and the education they will require: Evidence from occupational forecasts. *Educational Researcher, 20*(2), 11-20.

Bishop, J. H., & Carter, S. (1991b). The worsening shortage of college-graduate workers. *Educational Evaluation and Policy Analysis, 13*(3), 221-255.

Blien, U., & Tassinopoulos, A. (2001). Forecasting regional employment with the ENTROP method. *Regional Studies, 35*(2), 113-124.

Borghans, L., & Willems, E. (1998). Interpreting gaps in manpower forecasting models. *Labour, 12*(4), 633-41.

Byers, J. D., & Peel, D.A.. (1994). Non-linear adjustment of real wages, employment and output in the UK. *Applied Economics, 26*(5), 459-63.

Castley, R. (1996). Policy-focused approach to manpower planning. *International Journal of Manpower, 17*(3), 15-24.

Corvers, F., & Heijke, H. (2004). *Forecasting the labour market by occupations and education: Some key issues.* Retrieved 7/12, 2006, from http://www.roa.unimaas.nl/pdf%20publications/2004/ROA-W-2004_4.pdf

Fauvel, Y., Paquet, A., & Zimmerman, C. (1999). *Short-term forecasting of national and provincial employment in Canada* (Applied Research Branch of Strategic Policy No. R-99-6E). Hull, Quebec: Human Resources Development Canada.

Foot, D K., & Meltz, N. M. (1992). An ex post evaluation of Canadian occupational projections, 1961-1981. *Relations Industrielles, 47*(2), 268.

Heijke, H. (1996). *Labour market information for educational investments.* Retrieved 7/12, 2006, from http://www.roa.unimaas.nl/pdf%20publications/1996/roa-w-1996_2E.pdf

Hofler, R. A., & Spector, L. C. (1993). Employment determination in macroeconomic models: Some empirical evidence. *Journal of Macroeconomics, 15*(1), 123-38.

Kolb, R. A., & Stekler, H. O. (1992). Information content of long-term employment forecasts. *Applied Economics, 24*(6), 593-96.

Krolzig, H., Marcellino, M., & Mizon, G. E. (2002). A markov-switching vector equilibrium correction model of the UK labour market. *Empirical Economics, 27*(2), 233-54.

LeSage, J. P. (1990b). A comparison of the forecasting ability of ECM and VAR models. *Review of Economics and Statistics, 72*(4), 664-71.

LeSage, J. P. (1990a). Forecasting metropolitan employment using an export-base error-correction model. *Journal of Regional Science, 30*(3), 307-23.

Li, D. T., & Dorfman, J. H. (1995). A robust approach to predicting fluctuations in state-level employment growth. *Journal of Regional Science, 35*(3), 471-84.

London Economics. (2002). *Investigating the use of regional economic forecasting models in Wales.* Retrieved 7/12, 2006, from http://www.futureskillswales.com/uploads/documents/209.pdf

Longhi, S., & Nijkamp, P. (2005). *Forecasting regional labour market developments under spatial heterogeneity and spatial autocorrelation* (Discussion Paper No. TI 2005-041/3). Amsterdam, Netherlands: Tinbergen Institute.

Longhi, S., Nijkamp, P., Reggiani, A.., & Maierhofer, E. (2005). Neural network modeling as a tool for forecasting regional employment patterns. *International Regional Science Review, 28*(3), 330-46.

Magura, M. (1998). IO and spatial information as Bayesian priors in an employment forecasting model. *Annals of Regional Science, 32*(4), 495-503.

Partridge, M. D., & Rickman, D. S. (1998). Generalizing the Bayesian vector autoregression approach for regional interindustry employment forecasting. *Journal of Business and Economic Statistics, 16*(1), 62-72.

Patuelli, R., Longhi, S., Reggiani, A., & Nijkamp, P. (2003). Multicriteria analysis of neural network forecasting models: An application to German regional labour markets. *Studies in Regional Science, 33*(3), 205-29.

Puri, A., & Soydemir, G. (2000). Forecasting industrial employment figures in southern California: A Bayesian vector autoregressive model. *Annals of Regional Science, 34*(4), 503-14.

Rickman, D. S. (2001). Using input-output information for Bayesian forecasting of industry employment in a regional econometric model. *International Regional Science Review, 24*(2), 226-44.

Rickman, D. S., & Miller, S. R. (2002). An evaluation of alternative strategies for incorporating interindustry relationships into a regional employment forecasting model. *Review of Regional Studies, 32*(1), 133-47.

Rosenthal, N. H. (1999). The quality of BLS projections: A historical account. *Monthly Labor Review, 122*(5), 27-35.

Sarantis, N., & Swales, C. (1999). Modelling and forecasting regional service employment in Great Britain. *Economic Modeling*, 16(3), 429-53.

Smith, D. A.. (2002). Forecasting future skill needs in Canada; forecasting future skill needs in Canada. *Forecasting labour markets in OECD countries: Measuring and tackling mismatches,* 66-89.

Sweeney, S. H. (2004). Regional occupational employment projections: Modeling supply constraints in the direct-requirements approach*. *Journal of Regional Science, 44*(2), 263-288.

Trivez, F. J., & Mur, J. (1999). A short-term forecasting model for sectoral regional employment. *Annals of Regional Science, 33*(1), 69-91.

Turner, D. S., Wallis, K. F., & Whitley, J. D. (1994). Differences in the properties of large-scale macroeconometric models: The role of labour market specifications; differences in the properties of large-scale macroeconometric models: The role of labour market specifications. *Macroeconometric modelling.Volume 2.From the 1960s,* , 352-79.

van Eijs, P., & Borghans, L. (1996). The use of RAS in manpower forecasting: A microeconomic approach. *Economic Modelling, 13*(2), 257-87.

West, C. T. (2003). Structural regional factors that determine absolute and relative accuracy of U.S. regional labor market forecasts. *Journal of Agricultural and Applied Economics, 35*(0), 121-35.

Willems, E. (1996). *Manpower forecasting and modelling replacement demand: An overview*. Retrieved 7/12, 2006, from http://www.roa.unimaas.nl/pdf%20publications/1996/roa-w-1996_4E.pdf

Wong, J. M. W., Chan, A. P. C., & Chiang, Y. H. (2005). Time series forecasts of the construction labour market in Hong Kong: The Box-Jenkins approach. *Construction Management & Economics, 23*(9), 979-991.

Background Articles

Allen, P. G., & Morzuch, B. J. (2006). Twenty-five years of progress, problems, and conflicting evidence in econometric forecasting. what about the next 25 years? *International Journal of Forecasting, 22*(3), 475-492.

Alpert, A.., & Auyer, J. (2003). The 1988-2000 employment projections: How accurate were they? *Occupational Outlook Quarterly, 47*(1), 2-21.

Armstrong, J. S. (2006). Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting, 22*(3), 583-598.

Bishop, J. (1992). Is a skills shortage coming? *Workforce,* Spring, 15-30.

Bishop, J. H., & Carter, S. (1991a). How accurate are recent BLS occupational projections? *Monthly Labor Review, 114*(10), 37-43.

Bishop, J. H., & Carter, S. (1990). *The deskilling vs. upskilling debate: The role of BLS projections. working paper # 90-14*. U.S.; New York:

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting, 22*(3), 547-581.

Briscoe, G., & Wilson, R. (2003). Modelling UK occupational employment. *International Journal of Manpower, 24*(5), 568-589.

Burns, M., & Shanahan, M. (2000). *Labour market models and their use in projecting vocational eduaton and training requirements.* Retrieved 7/12, 2006, from http://www.ncver.edu.au/research/proj/nr7040.pdf

Campbell, C. P. (1997b). Labour market signalling approaches. *Journal of European Industrial Training, 21*(8/9), 284-290.

Campbell, C. P. (1997a). Workforce projection and forecasting approaches. *Journal of European Industrial Training, 21*(8/9), 281-283.

Centre d'étude sur l'emploi et la technologie. (2006). *Highly skilled workers in québec: Dynamic portrait of the labour market and adjustment factors.* Retrieved 7/12, 2006, from http://www.cetech.gouv.qc.ca/Site/Documents/Portrait%20THQ%20English-Overview.pdf

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting, 22*(3), 443-473.

Ellis, S. P. (2003). Anticipating employers' skills needs: The case for intervention. *International Journal of Manpower, 24*(1), 83-96.

Faberman, R. J. (2005). *Studying the labor market with the job openings and labor turnover survey* (Working Paper No. 388). Washington, DC: U.S. Department of Labor. (labor market search, Beveridge Curve)

Government of British Columbia: Ministry of Economic Development. (2006). *How do we forecast skills shortages?* Retrieved 07/12, 2006, from http://www.ecdev.gov.bc.ca/ProgramsAndServices/SkillsDevelopment/how-can-we-forecast.htm

Haskel, J., & Holt, R. (1999). *Anticipating future skill needs: Can it be done? does it need to be done?* Retrieved 07/12, 2006, from http://www.etechb.co.uk/reslib/Anticipating Future Skills Needs - Sept 1999.doc

He, C., & Xu, X. (2005). Combination of forecasts using self-organizing algorithms. *Journal of Forecasting, 24*(4), 269-78.

Hettinger, J. (1998). Where the BLS numbers "really" come from. *Techniques: Making Education and Career Connections, 73*(6), 36-39.

Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting, 22*(3), 493-518.

Ministry of Advanced Education: Planning and Data Management Branch. (2005). *Employment outlook for British Columbia: COPS BC unique scenario for 2003 to 2013*. Retrieved September 19, 2006, from http://www.aved.gov.bc.ca/labourmarketinfo/reports/COPS_BCUnique_paper.pdf

National Observatory of Vocational Training and Labour Market. (1999). *Forecasting education and training needs in transition economies: Lessons from the western European experience.* Retrieved 7/12, 2006, from http://www.nvf.cz/publikace/pdf_publikace/observator/eng/forecast_lessons.pdf

Neugart, M., & Schomann, K. (2002). Why forecast and for whom? Some introductory remarks. *Forecasting labour markets in OECD countries: Measuring and tackling mismatches,* , 1-25.

Psacharopoulos, G. (1991). From manpower planning to labour market analysis. *International Labour Review, 130*(4), 459-70.

Rosenthal, N. H. (1992). Evaluating the 1990 projections of occupational employment. *Monthly Labor Review, 115*(8), 32-48.

Rosenthal, N. H., Fullerton, H. N.,Jr, Andreassen, A.., & Veneri, C. M. (1997). Evaluating the 1995 BLS projections. *Monthly Labor Review, 120*(9), 3-31.

Schmidt, S. L., Schomann, K. & Tessaring, M. (2003). *Early identification of skill needs in Europe.* Retrieved 7/12, 2006, from http://www.trainingvillage.gr/etv/Upload/Projects_Networks/Skillsnet/Publications/3029_42.pdf

Statistics Canada. (2006). *Definitions, data sources, and methods: Occupation.* Retrieved Sept 20, 2006, from http://www.statcan.ca/english/concepts/definitions/work.htm

Van Adams, A., & Others. (1992). Market-based manpower planning with labour market signals. *International Labour Review, 131*(3), 261-279.

Veneri, C. M. (1999). Can occupational labor shortages be identified using available data? *Monthly Labor Review, 122*(3), 15-21.

Volterra Consulting Ltd. (2005). *Reviewing Future Skills Scotland's Forward Looking Approach.* Retrieved 7/12, 2006, from http://www.futureskillsscotland.org.uk/web/site/home/Reports/NationalReports/Report_Reviewing_Futureskills_Scotlands_Forward_Looking_Approach.asp